

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
28 September 2006 (28.09.2006)

PCT

(10) International Publication Number  
**WO 2006/102497 A2**

- (51) International Patent Classification: Not classified
- (21) International Application Number: PCT/US2006/010539
- (22) International Filing Date: 22 March 2006 (22.03.2006)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/664,550 22 March 2005 (22.03.2005) US
- (71) Applicant (for all designated States except US): **VASCULAR INFORMATION COMPANY** [US/US]; 750 Battery St., Suite 400, San Francisco, California 94111 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **TABIBIAZAR, Raymond** [US/US]; 2150 Sterling Avenue, Menlo Park, California 94025 (US). **QUERTERMOUS, Thomas** [US/US]; 810 Lathrop Drive, Stanford, California 94305 (US).
- (74) Agents: **JACOBSON, Jill, A.** et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, California 94304-1018 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR DIAGNOSIS, MONITORING AND DEVELOPMENT OF THERAPEUTICS FOR TREATMENT OF ATHEROSCLEROTIC DISEASE

(57) Abstract: Polynucleotide sequences are provided that correspond to genes that are differentially expressed in atherosclerotic disease conditions. Methods for using these sequences to detect gene expression and/or for transcriptional profiling in mammals are also provided. The polynucleotide sequences of the invention may be used, for example, to diagnose atherosclerotic disease, to monitor extent of progression or efficacy of treatment or to assess prognosis of atherosclerotic disease, and/or to identify compounds effective to treat an atherosclerotic disease condition.



WO 2006/102497 A2

## **METHODS AND COMPOSITIONS FOR DIAGNOSIS, MONITORING AND DEVELOPMENT OF THERAPEUTICS FOR TREATMENT OF ATHEROSCLEROTIC DISEASE**

### FIELD OF THE INVENTION

[0001] This application is in the field of atherosclerotic disease. In particular, this invention relates to methods and compositions for diagnosing, monitoring, and development of therapeutics for atherosclerotic disease.

### BACKGROUND OF THE INVENTION

[0002] Atherosclerosis is the primary cause of heart disease and stroke (Kannel and Belanger (1991) *Am. Heart J.* 121:951-57), and is the most common cause of morbidity and mortality in the United States (NHLBI Morbidity and Mortality Chartbook, National Heart, Lung, and Blood Institute, Bethesda, MD, May, 2002; NHLBI Fact Book, Fiscal Year 2003, pp. 35-53, National Heart, Lung, and Blood Institute, Bethesda, MD, February, 2004).

Atherosclerosis is currently conceptualized as a chronic inflammatory disease of the arterial vessel wall that develops due to complex interactions between the environment and the genetic makeup of an individual (Ross (1999) *N Engl J Med* 340:115-26). Development of an atherosclerotic plaque occurs in stages, beginning with simple fatty streak formation and culminating in complex calcified lesions containing abnormal accumulation of smooth muscle cells, inflammatory cells, lipids, and necrotic debris. It is likely that the various stages of atherosclerotic disease are governed by a set of genes that are expressed by a variety of cell types present in the vessel wall.

[0003] The propensity for developing atherosclerosis is dependent on underlying genetic risk, and varies as a function of age and exposure to environmental risk factors. However, despite the chronic nature of atherosclerotic disease, knowledge regarding temporal gene expression during the course of disease progression is very limited. The prolonged, chronic, and unpredictable nature of the disease in humans, by virtue of heterogeneous genetic and environment factors, has limited systematic temporal gene expression studies in humans.

[0004] The roles of a limited number of genes that are differentially expressed in vascular disease have been identified, and a few of these genes linked through mechanistic studies to

disease processes (Glass and Witztum (2001) *Cell* 104:503-16; Breslow (1996) *Science* 272:685-88; Lusa (2000) *Nature* 407:233-41). Recent efforts to identify disease related gene expression patterns have employed transcriptional profiling with DNA microarrays. However, these studies have included relatively small arrays (Wuttge et al. (2001) *Mol Med* 7:383-392) as well as limited time points, with the primary comparison between normal and late stage diseased tissue (Archacki et al. (2003) *Physiol Genomics* 15:65-74; Faber et al. (2002) *Curr Opin Lipidol* 13:545-552; McCaffrey et al. (2000) *J Clin Invest* 105:653-662; Randi et al. (2003) *J Throm Haemost* 1:829-835; Seo et al. (2004) *Arterioscler Thromb Vasc Biol* 24:1922-1927; Zohlhofer et al. (2001) *Mol Cell* 7:1059-1069. Utilizing microarrays in animal models, where a disease process can be studied over time, the impact of individual risk factors and perturbations on the expression of individual genes during disease development can be studied systematically without *a priori* knowledge of gene identity. The temporal expression patterns of the genes can then be correlated with the well-described disease stages.

**[0005]** There is a need for a comprehensive list of atherosclerosis-related genes that are predictive of atherosclerotic disease conditions, for use as diagnostic markers and for discovery of biochemical pathways involved in development of atherosclerotic disease and discovery and/or testing of new therapeutics.

#### BRIEF SUMMARY OF THE INVENTION

**[0006]** This invention provides compositions, methods, and kits for detection of gene expression, diagnosis, monitoring, and development of therapeutics with respect to atherosclerotic disease.

**[0007]** In one aspect, the invention provides a system for detecting gene expression, comprising at least two isolated polynucleotide molecules, wherein each isolated polynucleotide molecule detects an expressed gene product from a gene that is differentially expressed in atherosclerotic disease in a mammal. In one embodiment, the differentially expressed gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the differentially expressed gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in

SEQ ID NOs: 1-927. In various embodiments, a system for detecting gene expression comprises any of at least 3, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, or 100 of the isolated polynucleotide molecules described herein or their polynucleotide complements, or human homologs or orthologs thereof. In one embodiment, the gene expression system comprises at least two isolated polynucleotide molecules, wherein each isolated polynucleotide molecule detects an expressed gene product, wherein the gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927, wherein the gene is differentially expressed in atherosclerotic disease in a mammal, and wherein the gene expression system comprises at least 1, 3, 5, 10, 15, 20, 25, or 30 isolated polynucleotide molecules that detect genes corresponding to the polynucleotide sequences selected from the group consisting of SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.

**[0008]** In some embodiments, the isolated polynucleotide molecules are immobilized on an array, which may be selected from the group consisting of a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, a polynucleotide array, a cDNA array, a microtiter plate, a membrane, and a chip. The isolated polynucleotide molecules may be selected from the group consisting of synthetic DNA, genomic DNA, cDNA, RNA, or PNA. A gene corresponding to an isolated polynucleotide molecules described herein may be differentially expressed in any blood vessel or portion thereof which has developed an atherosclerotic or inflammatory disease, for example, the aorta, a coronary artery, the carotid artery, or a blood vessel of the peripheral vasculature.

**[0009]** In another aspect, the invention provides a kit comprising a system for detecting gene expression as described above. In one embodiment, the kit comprises an array comprising a system for detecting gene expression as described above.

**[0010]** In another aspect, the invention provides a method of detecting gene expression, comprising contacting products of gene expression with the system for detecting gene expression as described above. In one embodiment, the method comprises isolating mRNA, for example from a sample from individual who has or who is suspected of having an atherosclerotic disease, and hybridizing the RNA to the polynucleotide molecules from the system for detecting gene expression. In another embodiment, the method comprises isolating

mRNA, converting the RNA to nucleic acid derived from the RNA, *e.g.*, cDNA, and hybridizing the nucleic acid derived from the RNA to the polynucleotide molecules of the system for detecting gene expression. Optionally, the RNA may be amplified prior to hybridization to the system for gene expression. Optionally, the RNA is detectably labeled, and determination of presence, absence, or amount of an RNA molecule corresponding to a gene detected by a polynucleotide molecule of the system for detecting gene expression comprises detection of the label.

**[0011]** In another embodiment, the method for detecting gene expression comprises isolating proteins from an individual who has or who is suspected of having an atherosclerotic disease, and detecting the presence, absence, or amount of one or more proteins corresponding to the gene expression product of a gene that is differentially expressed in atherosclerotic disease and corresponds to a polynucleotide molecule of the system for detecting gene expression as described above. Detection may be via an antibody that recognizes the protein, for example, by contacting the isolated proteins with an antibody array.

**[0012]** In another aspect, the invention provides a method for diagnosing an atherosclerotic disease in an individual, comprising contacting polynucleotides derived from a sample from the individual with a system for detecting gene expression as described above. In one embodiment, the method comprises detecting hybridization complexes formed, if any, wherein presence, absence or amount of hybridization complexes formed from at least one of the polynucleotides from the individual is indicative of presence or absence of the atherosclerotic disease. In another embodiment, the method comprises comparing levels of expression of the genes with a molecular signature indicative of the presence or absence of the atherosclerotic disease.

**[0013]** In another aspect, the invention provides a method for assessing extent of progression of atherosclerotic disease in an individual, comprising contacting polynucleotides derived from a sample from the individual with a system for detecting gene expression as described above. In one embodiment, the method comprises detecting hybridization complexes formed, if any, wherein presence, absence or amount of hybridization complexes formed from at least one of the polynucleotides from the individual is indicative of extent of progression of the atherosclerotic disease. In another embodiment, the method comprises detecting hybridization complexes formed, if any, and comparing levels of expression of the genes with a molecular signature indicative of extent of progression of the atherosclerotic disease.

**[0014]** In another aspect, the invention provides a method of assessing efficacy of treatment of atherosclerotic disease in an individual, comprising contacting polynucleotides derived from a sample from the individual with a system for detecting gene expression as described above. In one embodiment, the method comprises detecting hybridization complexes formed, if any, wherein presence, absence or amount of hybridization complexes formed from at least one of the polynucleotides from the individual is indicative of extent of progression of the atherosclerotic disease. In another embodiment, the method comprises comparing levels of expression of the genes with a molecular signature indicative of extent of progression of the atherosclerotic disease.

**[0015]** In another aspect, the invention provides a method for determining prognosis of atherosclerotic disease in an individual, comprising contacting polynucleotides derived from a sample from the individual with a system for detecting gene expression as described above. In one embodiment, the method comprises detecting hybridization complexes formed, if any, wherein presence, absence or amount of hybridization complexes formed from at least one of the polynucleotides from the individual is indicative of prognosis of the atherosclerotic disease. In another embodiment, the method comprises comparing levels of expression of the genes with a molecular signature indicative of prognosis of the atherosclerotic disease.

**[0016]** In another aspect, the invention provides a method for identifying a compound effective to treat an atherosclerotic disease, comprising administering a test compound to a mammal with an atherosclerotic disease condition and contacting polynucleotides derived from a sample from the mammal with a system for detecting gene expression as described above. In one embodiment, the method comprises detecting hybridization complexes formed, if any, wherein presence, absence or amount of hybridization complexes formed from at least one of the polynucleotides from the individual is indicative of treatment of the disease. In another embodiment, the invention comprises detecting hybridization complexes formed, if any, and comparing levels of expression of the genes with a molecular signature indicative of treatment of the disease.

**[0017]** In another aspect, the invention provides a method of monitoring atherosclerotic disease in a mammal, comprising detecting the expression level of at least one, at least two, at least ten, at least one hundred, or more genes selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927. In some embodiments, at least

one of the genes for which expression level is detected is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In one embodiment, the atherosclerotic disease comprises coronary artery disease. In one embodiment, the atherosclerotic disease comprises carotid atherosclerosis. In one embodiment, the atherosclerotic disease comprises peripheral vascular disease. In some embodiments, the expression level of said gene(s) is detected by measuring the RNA expression level. In one embodiment, RNA is isolated from the individual prior to detection of the RNA expression level. Measurement of RNA expression level may comprise amplifying RNA from an individual, for example, by polymerase chain reaction (PCR), using a primer that is complementary to a polynucleotide sequence corresponding to a gene to be detected, wherein the gene corresponds to a polynucleotide sequence selected from the group of genes depicted in SEQ ID NOs: 1-927. In some embodiments, a primer is used that is complementary to a polynucleotide sequence corresponding to a gene to be detected, wherein the gene corresponds to a polynucleotide sequence selected from the group of genes depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. Measurement of RNA expression level may comprise hybridization of RNA from the individual to a polynucleotide corresponding to a gene to be detected, wherein the gene corresponds to a polynucleotide sequence selected from the group of genes depicted in SEQ ID NOs: 1-927. In some embodiments, RNA from the individual is hybridized to a polynucleotide corresponding to a gene to be detected, wherein the gene to be detected is selected from the group of genes depicted in 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In some embodiments, gene expression level is detected by measuring the expressed protein level. In some embodiments, the method further comprises selecting an appropriate therapy for treatment or prevention of the atherosclerotic disease. In some embodiments, gene expression level, for example, RNA or protein level, is detected in serum from an individual.

[0018] In another aspect, the invention provides a method of monitoring atherosclerotic disease in an individual, comprising detecting RNA expressed from at least one gene selected from the group of genes corresponding to at least one polynucleotide sequence depicted in SEQ ID NOs:1-927. In one embodiment, the at least one gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In one embodiment, the method comprises measuring the expressed RNA in serum from the individual.

[0019] In another aspect, the invention provides a method of monitoring atherosclerotic disease in an individual, comprising detecting protein expressed from at least one gene selected from the group of genes corresponding to at least one polynucleotide sequence depicted in SEQ ID NOs:1-927. In one embodiment, the at least one gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In one embodiment, the method comprises measuring the expressed protein in serum from the individual.

#### BRIEF DESCRIPTION OF THE FIGURES

[0020] **Figure 1** depicts the experimental design of the experiments described in Example 1. ApoE deficient mice (C57BL/6J-*ApoE*<sup>5mlUnc</sup>), were fed non-cholesterol-containing high-fat diet from 4 weeks of age for a maximum period of 40 weeks. Aortas were obtained for transcriptional profiling at pre-determined time intervals corresponding to various stages of atherosclerotic plaque formation. For each time point, aortas from 15 mice were combined into 3 pools for microarray replicate studies. To eliminate gene expression differences due to aging, diet, and genetic differences, a number of control groups were also used at each time point, including *apoE* deficient mice on normal chow, as well as C57BL/6 and C3H/HeJ wild type mice on both normal and atherogenic diets.

[0021] **Figure 2** depicts quantification of atherosclerotic disease in the experiments described in Example 1. Percent lesion area was determined by calculating the ratio of atherosclerotic area versus total surface area of the aorta. ApoE-deficient mice (n=7) on high-fat diet were compared to other control mice (n=5-7 for each mouse/diet combination). Representative time intervals were used for analysis, including baseline (T00) measurements in mice prior to initiation of diet at 4 weeks of age and end point measurements corresponding to 40 weeks (T40) on either high-fat or normal diet. At T00, there were no statistically significant differences in lesion area among the various conditions. At 40 weeks on high-fat diet, the controls did not develop any lesions. In contrast to the control mice, the ApoE-deficient mice on normal chow and on high-fat diet had significantly larger atherosclerotic area (14.00% +/- 3.92%,  $p < 0.0001$ , and 37.98% +/- 6.3%,  $p < 0.0001$ , respectively.)

[0022] **Figure 3** depicts atherosclerosis genes identified in the experiments described in Example 1. Employing a newly-developed statistical algorithm which relies on permutation analysis and generalized regression, atherosclerosis-related genes were identified. Selecting the genes on the basis of their false detection rate ( $FDR < 0.05$ ) and depicting their expression with a heatmap (ordered by hierarchical clustering), demonstrates profiles which closely correlate with disease progression. The heatmap is a graphic representation of expression patterns of 6 parallel time course studies with time progressing from left to right for each of the 6 sets of strain-diet combination. Each set of the strain-diet combination therefore contains 15 columns (3 for each of 5 time points). Each row represents the row normalized expression pattern of a single gene. The dominant temporal pattern of expression is one that increases linearly with time (667 genes). Fewer genes (64) reveal an opposite pattern. HF: high-fat diet; NC: normal chow.

[0023] **Figure 4** depicts time-related patterns of gene expression in atherosclerosis observed in the experiments described in Example 1. Using AUC analysis, a number of distinct time-related patterns of gene expression in ApoE-deficient mice on high-fat diet were observed. Eight different time-related patterns are depicted, with the y-axis representing normalized gene expression values and the x-axis representing 6 different time points from time 0 to 40 weeks. The genes in each pattern were clustered based on positive correlation values. The mean distance of genes from the center of each cluster is noted in parentheses for each pattern. Using

enrichment analysis for each cluster of genes, specific pathways were found to be associated with these patterns that reflect particular biological processes.

[0024] **Figure 5** depicts the identification and validation of mouse atherosclerotic disease classifier genes as determined in the experiments described in Example 1. **Figure 5A** depicts identification of the classification gene set. The SVM algorithm described in Example 1 was employed to rank genes based on their abilities to accurately discriminate between 5 time points in ApoE-deficient mice on high-fat diet. An optimal set of 38 genes was identified to classify the experiments at a minimal error rate of 15%. The optimal 15% error rate was determined with a 1000 step cross-validation method with 25% of the experiments employed as the test group and the rest as the training group. **Figure 5B** depicts classification of an independent mouse atherosclerosis data set. Aortas of ApoE-deficient mice aged 16 weeks were used for gene expression profiling utilizing a different microarray and labeling protocol than in the experiment depicted in Figure 5A. Using the SVM algorithm, where known experiments were the five time points in the original experimental design and the independent set of experiments was the test set, these mice most closely classified with the 24 week time point. SVM scores for each experiment based on one-versus-all comparisons are represented graphically in a heatmap.

[0025] **Figure 6** depicts expression of atherosclerosis-related genes in human coronary artery disease, as described in Example 1. To investigate the expression profile of differently regulated mouse genes in human coronary artery atherosclerosis, 40 coronary artery samples with and without atherosclerotic lesions were used for transcriptional profiling. Atherosclerosis-associated mouse genes were matched to human orthologs/homologs by gene symbol and by known homology, and their expression was compared in human atherosclerotic plaques classified as lesion versus no lesion (SAM FDR<0.025). The expression of the top genes is represented graphically as a heatmap, where rows represent row normalized expression of each gene and the columns represent coronary artery samples. Calculated SAM FDR<0.009 for d-score 4.25-2.45, FDR<0.015 for d-score 2.41-2.357, FDR<0.025 for d-score 2.33-2.05.

[0026] **Figure 7** depicts the experimental design of the experiments described in Example 2. **Fig. 7A:** Four-week-old female C3H/HeJ (C3H) and C57Bl6 (C57) mice were fed normal chow vs. high-fat diet for the maximum period of 40 weeks. Triplicate microarray experiments were performed for each time point using 3 pools of 5 aortas at 0, 4, 10, 24, and 40 weeks on

either diet (total of 15 mice per time point). **Fig. 7B:** Data analysis overview. Of the 20,283 genes present on the array, 311 genes were found to be significantly differentially expressed between C3H and C57 mice at baseline (SAM FDR 10% and >1.5-fold change). Differential gene expression during aging was determined by comparing C57 vs. C3H time-course differences on normal and atherogenic high-fat diets using AUC analysis.

[0027] **Figure 8** depicts differential gene expression between C3H and C57 mice at baseline. The SAM analysis shown was associated with an FDR of 10%, and a total of 311 probes were identified as differentially regulated at this level of confidence. Lists represent a select group of genes (expressed sequence tags excluded) with higher expression in C3H (top 20 ranking genes) and C57 (top 45 ranking genes). The heatmap reflects normalized gene expression ratios and is organized with individual hybridizations for each of the 3 replicates for each mouse strain arranged along the x axis.

[0028] **Figure 9** depicts differential gene expression between C3H and C57 mice in response to normal aging. **Fig. 9A:** Response to aging was determined by comparing C57 vs. C3H time-course differences on normal diet (AUC analysis F statistic>10). **Fig. 9B:** Functional annotation of the 413 differentially expressed genes reveals differences in various biological processes, including growth and differentiation. The probability rates provided area based on Fisher exact test ( $P<0.02$ ). **Fig. 9C:** K-means clustering of the 413 genes reveals several profiles of gene expression. Clusters 1, 4, and 9 reveal increased gene expression in C3H vs. C57 mice, whereas clusters 2, 6, and 14 reveal the opposite pattern.

[0029] **Figure 10** depicts differential gene expression between C3H and C57 mice in response to high-fat diet. **Fig. 10A:** Response to atherogenic stimulus was determined by comparing C57 vs. C3H time-course differences on high –fat diet (AUC analysis F statistic>10). **Fig. 10B:** Functional annotation of the 509 differentially expressed genes reveals differences in various biological processes and cellular components. The probability rates provided are based on Fisher exact test ( $P<0.02$ ). **Fig. 10C:** K-means clustering of the 509 differentially expressed genes revealed several patterns of gene expression with clusters 3 and 9 exhibiting increased gene expression in C3H vs. C57 mice and clusters 8 and 10 with the opposite pattern.

[0030] **Figure 11** shows the results of evaluation in the apoE knockout model of genes identified as differentially expressed between C3H and C57 strains. **Fig. 11A:** ApoE knockout

mice (C57BL/6J-*ApoE*<sup>tm1Unc</sup>) were fed normal chow versus high-fat diet for the maximum period of 40 weeks. Triplicate microarray experiments were performed for each time point using 3 pools of 5 aortas at 0, 4, 10, 24, and 40 weeks for regular and high-fat diet groups (total of 15 mice per time point). SOMs were used to visualize patterns of expression of genes of interest. Genes which were differentially regulated by aging (**Fig. 9**, K-means clusters 1, 4, and 9 with higher expression in C3H and clusters 4, 6, and 14 with higher expression in C57) and genes identified with atherogenic stimuli (**Fig. 10**, K-means clusters 3 and 9 with higher expression in C3H and clusters 8 and 10 with opposite pattern) as well as genes which were differentially expressed at the baseline time point (**Fig. 8**), were grouped and their expression was studied using SOM analysis. SOM analysis reveals diverse patterns of expression of these genes throughout the development of atherosclerosis in apoE knockout mice. Cluster 8 contains genes that are consistently increasing in expression with progression of atherosclerosis. Pie charts reflect the analysis group from which the genes populating each cluster were derived. The relative size of sectors of the pie chart indicates the relative number of genes that are derived from the various staging groups. **Fig. 11B** lists genes with higher expression in C57 mice at baseline and in C3H mice at baseline or on a high fat diet.

#### DETAILED DESCRIPTION OF THE INVENTION

[0031] The invention provides polynucleotide sequences that correspond to genes that are differentially expressed in atherosclerotic disease conditions, and methods for using these sequences to detect gene expression and/or for transcriptional profiling in mammals. The polynucleotide sequences provided herein may be used, for example, to diagnose, assess extent of progression, assess efficacy of treatment of, to determine prognosis of, and/or to identify compounds effective to treat an atherosclerotic disease condition. The polynucleotide sequences herein may also be used in methods for elucidation of biochemical pathways that are involved in development and/or maintenance of atherosclerotic disease conditions.

#### ***General Techniques***

[0032] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, and biochemistry, which are within the skill of the art. Such techniques are explained fully in the literature, such as: *Molecular Cloning: A Laboratory Manual*, vol. 1-3, third edition (Sambrook et al., 2001); *Oligonucleotide Synthesis* (M.J. Gait, ed., 1984); *Methods in Enzymology* (Academic Press, Inc.); *Current Protocols in Molecular Biology* (F.M. Ausubel et al., eds., 1987); *PCR Cloning Protocols*, (Yuan and Janes, eds., 2002, Humana Press).

[0033] In addition to the above references, protocols for *in vitro* amplification techniques, such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Q $\beta$ -replicase amplification, and other RNA polymerase mediated techniques (*e.g.*, NASBA), useful, *e.g.*, for amplifying oligonucleotide probes of the invention, are found in Mullis et al., U.S. Patent No. (1987) 4,683,202; *PCR Protocols: A Guide to Methods and Applications* (Innis et al., eds.) Academic Press, Inc., San Diego, CA (1990); Arnheim and Levinson (1990) *C&EN* 36; *The Journal of NIH Research* (1991) 3:81; Kwoh et al. (1989) *Proc Natl Acad Sci USA* 86:1173; Guatelli et al. (1990) *Proc Natl Acad Sci USA* 87:1874; Lomell et al. (1989) *J Clin Chem* 35:1826; Landegren et al. (1988) *Science* 241:1077; Van Brunt (1990) *Biotechnology* 8:291; Wu and Wallace (1989) *Gene* 4:560; Barringer et al. (1990) *Gene* 89:117; Sooknanan and Malek (1995) *Biotechnology* 13:563. Additional methods, useful for cloning nucleic acids, include Wallace et al., U.S. Patent No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) *Nature* 369:684, and the references therein.

### **Definitions**

[0034] Unless defined otherwise, all scientific and technical terms are understood to have the same meaning as commonly used in the art to which they pertain. For the purpose of the present invention, the following terms are defined below.

[0035] As used herein, the term “gene expression system” or “system for detecting gene expression” refers to any system, device or means to detect gene expression and includes candidate libraries, oligonucleotide sets or probe sets.

[0036] The term “diagnostic oligonucleotide set” generally refers to a set of two or more oligonucleotides that, when evaluated for differential expression of their products, collectively yields predictive data. Such predictive data typically relates to diagnosis, prognosis, monitoring of therapeutic outcomes, and the like. In general, the components of a diagnostic oligonucleotide set are distinguished from nucleotide sequences that are evaluated by analysis of the DNA to directly determine the genotype of an individual as it correlates with a specified trait or phenotype, such as a disease, in that it is the pattern of expression of the components of the diagnostic nucleotide set, rather than mutation or polymorphism of the DNA sequence that provides predictive value. It will be understood that a particular component (or member) of a diagnostic nucleotide set can, in some cases, also present one or more mutations, or polymorphisms that are amenable to direct genotyping by any of a variety of well known analysis methods, e.g., Southern blotting, RFLP, AFLP, SSCP, SNP, and the like.

[0037] A “disease specific target oligonucleotide sequence” is a gene or other oligonucleotide that encodes a polypeptide, most typically a protein, or a subunit of a multi-subunit protein, that is a therapeutic target for a disease, or group of diseases.

[0038] A “candidate library” or a “candidate oligonucleotide library” refers to a collection of oligonucleotide sequences (or gene sequences) that by one or more criteria have an increased probability of being associated with a particular disease or group of diseases. The criteria can be, for example, a differential expression pattern in a disease state, tissue specific expression as reported in a sequence database, differential expression in a tissue or cell type of interest, or the like. Typically, a candidate library has at least 2 members or components; more typically, the library has in excess of about 10, or about 100, or about 500, or even more, members or components.

[0039] The term “disease criterion” is used herein to designate an indicator of a disease, such as a diagnostic factor, a prognostic factor, a factor indicated by a medical or family history, a genetic factor, or a symptom, as well as an overt or confirmed diagnosis of a disease associated with several indicators. A disease criterion includes data describing a patient’s health status, including retrospective or prospective health data, e.g., in the form of the patient’s medical history, laboratory test results, diagnostic test results, clinical events, medications, lists, response(s) to treatment and risk factors, etc.

[0040] The terms “molecular signature” or “expression profile” refers to the collection of expression values for a plurality (*e.g.*, at least 2, but frequently at least about 10, about 30, about 100, about 500, or more) of members of a candidate library. In many cases, the molecular signature represents the expression pattern for all of the nucleotide sequences in a library or array of candidate or diagnostic nucleotide sequences or genes. Alternatively, the molecular signature represents the expression pattern for one or more subsets of the candidate library.

[0041] The terms “oligonucleotide” and “polynucleotide” and “nucleic acid,” used interchangeably herein, refer to a polymeric form of two or more nucleotides of any length and any three-dimensional structure (*e.g.*, single-stranded, double-stranded, triple-helical, etc.), which contain deoxyribonucleotides, ribonucleotides, and/or analogs or modified forms of deoxyribonucleotides or ribonucleotides.. Nucleotides may be DNA or RNA, and may be naturally occurring, or synthetic, or non-naturally occurring. A nucleic acid of the present invention may contain phosphodiester bonds or an alternate backbone, comprising, for example, phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidite linkages, and peptide nucleic acid backbones and linkages. The term polynucleotide includes peptide nucleic acids (PNA).

[0042] The terms “polypeptide,” “peptide,” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The term also includes variants on the traditional peptide linkage joining the amino acids making up the polypeptide.

[0043] An “isolated” or “purified” polynucleotide or polypeptide is one that is substantially free of the materials with which it is associated in nature. By substantially free is meant at least 50%, preferably at least 70%, more preferably at least 80%, and even more preferably at least 90% free of the materials with which it is associated in nature.

[0044] As used herein, “individual” refers to a vertebrate, typically a mammal, such as a human, a nonhuman primate, an experimental animal, such as a mouse or rat, a pet animal, such as a cat or dog, or a farm animal, such as a horse, sheep, cow, or pig.

[0045] The term “healthy individual,” as used herein, is relative to a specified disease or disease criterion, *e.g.*, the individual does not exhibit the specified disease criterion or is not

diagnosed with the specified disease. It will be understood that the individual in question can exhibit symptoms, or possess various indicator factors, for another disease.

[0046] Similarly, an “individual diagnosed with a disease” refers to an individual diagnosed with a specified disease (or disease criterion). Such an individual may, or may not, also exhibit a disease criterion associated with, or be diagnosed with another (related or unrelated) disease.

[0047] An “array” is a spatially or logically organized collection, *e.g.*, of oligonucleotide sequences or nucleotide sequence products such as RNA or proteins encoded by an oligonucleotide sequence. In some embodiments, an array includes antibodies or other binding reagents specific for products of a candidate library.

[0048] When referring to a pattern of expression, a “qualitative” difference in gene expression refers to a difference that is not assigned a relative value. That is, such a difference is designated by an “all or nothing” valuation. Such an all or nothing variation can be, for example, expression above or below a threshold of detection (an on/off pattern of expression). Alternatively, a qualitative difference can refer to expression of different types of expression products, *e.g.*, different alleles (*e.g.*, a mutant or polymorphic allele), variants (including sequence variants as well as post-translationally modified variants), etc.

[0049] In contrast, a “quantitative” difference, when referring to a pattern of gene expression, refers to a difference in expression that can be assigned a numerical value, such as a value on a graduated scale, (*e.g.*, a 0-5 or 1-10 scale, a + - +++ scale, a grade 1- grade 5 scale, or the like; it will be understood that the numbers selected for illustration are entirely arbitrary and in no-way are meant to be interpreted to limit the invention).

[0050] The term “monitoring” is used herein to describe the use of gene sets to provide useful information about an individual or an individual’s health or disease status. “Monitoring” can include, for example, determination of prognosis, risk-stratification, selection of drug therapy, assessment of ongoing drug therapy, determination of effectiveness of treatment, prediction of outcomes, determination of response to therapy, diagnosis of a disease or disease complication, following of progression of a disease or providing any information relating to a patient’s health status over time, selecting patients most likely to benefit from experimental therapies with known molecular mechanisms of action, selecting patients most likely to benefit from approved drugs with known molecular mechanisms where that mechanism may be important in a small subset of a disease for which the medication may not have a label,

screening a patient population to help decide on a more invasive/expensive test, for example, a cascade of tests from a non-invasive blood test to a more invasive option such as biopsy, or testing to assess side effects of drugs used to treat another indication.

### ***System for detecting gene expression***

[0051] The invention provides a system for detecting expression of genes that are differentially expressed in atherosclerotic disease. In one embodiment, the system for detecting gene expression detects at least two expressed gene products of genes selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the system for detecting gene expression detects at least two expressed gene products of genes selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927. The term “corresponding” as used herein in the context of a gene corresponding to a polynucleotide sequence depicted in the Sequence Listing refers to a gene that is detectable by interaction of a product of expression of the gene (*e.g.*, mRNA, protein) or a product derived from a product of expression of the gene (*e.g.*, cDNA) with the system for detecting gene expression. The polynucleotide sequences represented by Sequence Identification Nos. 1-927 and accompanying identifying information are depicted in Table 1 below. These sequences have been shown to be differentially expressed in atherosclerosis in mice (see Example 1). The 60mer sequences represented in Table 1 are encompassed within the genes indicated therein. The gene sequences are obtainable from publicly available databases such as GenBank, and at <http://www.ncbi.nlm.nih.gov> or <http://source.stanford.edu/cgi-bin/source/sourceSearch>, using the identifying information provided in Table 1.

[0052] In one embodiment, the system for detecting gene expression includes at least two isolated polynucleotide molecules, each of which detects an expressed gene product of a gene that is differentially expressed in atherosclerotic disease in a mammal. The gene expression system includes at least two isolated polynucleotides that each comprise at least a portion of a sequence depicted in the Sequence Listing or its complement (*i.e.*, a polynucleotide sequence capable of hybridizing to a sequence depicted in the sequence listing). A system for detecting

gene expression in accordance with the invention may include any of at least 2, 3, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, or 100 polynucleotides each comprising at least a portion of a polynucleotide depicted in the Sequence Listing or a polynucleotide complement thereof.

**[0053]** It is understood that the polynucleotides of the invention may have slightly different sequences than those identified herein. Such sequence variations are understood to those of ordinary skill in the art to be variations in the sequence that do not significantly affect the ability of the sequences to detect gene expression. For example, homologs and variants of the polynucleotides disclosed herein may be used in the present invention. Homologs and variants of these polynucleotide molecules possess a relatively high degree of sequence identity when aligned using standard methods. Polynucleotide sequences encompassed by the invention have at least 40-50, 50-60, 70-80, 80-85, 85-90, 90-95 or 95-100% sequence identity to the sequences disclosed herein.

**[0054]** It is understood that for expression profiling, variations in the disclosed polynucleotide sequences will still permit detection of gene expression. The degree of sequence identity required to detect gene expression varies depending on the length of an oligonucleotide. For example, for a 60mer (*i.e.*, an oligonucleotide with 60 nucleotides), 6-8 random mutations or 6-8 random deletions do not affect gene expression detection. Hughes, T.R., et al. (2001) *Nature Biotechnology* 19:343-347. As the length of the polynucleotide sequence is increased, the number of mutations or deletions permitted while still allowing gene expression detection is increased.

**[0055]** As will be appreciated by those skilled in the art, the sequences of the present invention may contain sequencing errors. For example, there may be incorrect nucleotides, frameshifts, unknown nucleotides, or other types of sequencing errors in any of the sequences; however, the correct sequences will fall within the homology and stringency definitions herein.

**[0056]** In some embodiments, polynucleotide molecules are less than about any of the following lengths (in bases or base pairs): 10,000; 5000; 2500; 2000; 1500; 1250; 1000; 750; 500; 300; 250; 200; 175; 150; 125; 100; 75; 50; 25; 10. In some embodiments, polynucleotide molecules are greater than about any of the following lengths (in bases or base pairs): 10; 15; 20; 25; 30; 40; 50; 60; 75; 100; 125; 150; 175; 200; 250; 300; 350; 400; 500; 750; 1000; 2000; 5000; 7500; 10,000; 20,000; 50,000. Alternately, a polynucleotide molecule can be any of a range of sizes having an upper limit of 10,000; 5000; 2500; 2000; 1500; 1250; 1000; 750; 500;

300; 250; 200; 175; 150; 125; 100; 75; 50; 25; or 10 and an independently selected lower limit of 10; 15; 20; 25; 30; 40; 50; 60; 75; 100; 125; 150; 175; 200; 250; 300; 350; 400; 500; 750; 1000; 2000; 5000; or 7500, wherein the lower limit is less than the upper limit.

[0057] The isolated polynucleotides of the system for detecting gene expression may include DNA or RNA or a combination thereof, and/or modified forms thereof, and/or may also include a modified polynucleotide backbone. In some embodiments, the isolated polynucleotides are selected from the group consisting of synthetic oligonucleotides, genomic DNA, cDNA, RNA, or PNA.

[0058] In one embodiment, the system for detecting gene expression comprises two antibody molecules or antigen binding fragments thereof, each of which detects an expressed gene product (*e.g.*, a polypeptide) of a gene that is differentially expressed in atherosclerotic disease in a mammal.

[0059] As used herein, "atherosclerotic disease" refers to a vascular inflammatory disease characterized by the deposition of atheromatous plaques containing cholesterol, lipids, and inflammatory cells within the walls of large and medium-sized blood vessels, which can lead to hardening of blood vessels, stenosis, and thrombotic and embolic events. Atherosclerosis includes coronary vascular disease, cerebral vascular disease, and peripheral vascular disease. The term "atherosclerotic disease" as used herein includes any condition associated with atherosclerosis in a mammal in which differential gene expression may be detected by a system for detecting gene expression as described herein. Examples of such atherosclerotic disease conditions include, but are not limited to, coronary artery disease (*e.g.*, stable angina, unstable angina, exertional angina, myocardial infarction, congestive heart failure, sudden cardiac death, atrial fibrillation), cerebral vascular disease (*e.g.*, stroke, cerebrovascular accident (CVA), transient ischemic attack (TIA), cerebral infarction, cerebral intermittent claudication), peripheral vascular disease (*e.g.*, claudications), extracranial carotid disease, carotid plaque, and carotid bruit.

### ***Arrays***

[0060] In some embodiments, a system for detecting gene expression in accordance with the invention is in the form of an array. "Microarray" and "array," as used interchangeably herein, comprise a surface with an array, preferably ordered array, of putative binding (*e.g.*, by

hybridization) sites for a biochemical sample (target) which often has undetermined characteristics. In one embodiment, a microarray refers to an assembly of distinct polynucleotide or oligonucleotide probes immobilized at defined positions on a substrate. Arrays may be formed on substrates fabricated with materials such as paper, glass, plastic (*e.g.*, polypropylene, nylon, polystyrene), polyacrylamide, nitrocellulose, silicon, optical fiber or any other suitable solid or semi-solid support, and configured in a planar (*e.g.*, glass plates, silicon chips) or three-dimensional (*e.g.*, pins, fibers, beads, particles, microtiter wells, capillaries) configuration. Probes forming the arrays may be attached to the substrate by any number of ways including (i) *in situ* synthesis (*e.g.*, high-density oligonucleotide arrays) using photolithographic techniques (see, Fodor et al., *Science* (1991), 251:767-773; Pease et al., *Proc. Natl. Acad. Sci. U.S.A.* (1994), 91:5022-5026; Lockhart et al., *Nature Biotechnology* (1996), 14:1675; U.S. Pat. Nos. 5,578,832; 5,556,752; and 5,510,270); (ii) spotting/printing at medium to low-density (*e.g.*, cDNA probes) on glass, nylon or nitrocellulose (Schena et al., *Science* (1995), 270:467-470, DeRisi et al., *Nature Genetics* (1996), 14:457-460; Shalon et al., *Genome Res.* (1996), 6:639-645; and Schena et al., *Proc. Natl. Acad. Sci. U.S.A.* (1995), 93:10539-11286); (iii) by masking (Maskos and Southern, *Nuc. Acids. Res.* (1992), 20:1679-1684) and (iv) by dot-blotting on a nylon or nitrocellulose hybridization membrane (see, *e.g.*, Sambrook et al., Eds., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Vol. 1-3, Cold Spring Harbor Laboratory (Cold Spring Harbor, N.Y.)). Probes may also be noncovalently immobilized on the substrate by hybridization to anchors, by means of magnetic beads, or in a fluid phase such as in microtiter wells or capillaries. The probe molecules are generally nucleic acids such as DNA, RNA, PNA, and cDNA but may also include proteins, polypeptides, oligosaccharides, cells, tissues and any permutations thereof which can specifically bind the target molecules.

**[0061]** For example, microarrays, in which either defined cDNAs or oligonucleotides are immobilized at discrete locations on, for example, solid or semi-solid substrates, or on defined particles, enable the detection and/or quantification of the expression of a multitude of genes in a given specimen.

**[0062]** Several techniques are well-known in the art for attaching nucleic acids to a solid substrate such as a glass slide. One method is to incorporate modified bases or analogs that contain a moiety that is capable of attachment to a solid substrate, such as an amine group, a

derivative of an amine group or another group with a positive charge, into the amplified nucleic acids. The amplified product is then contacted with a solid substrate, such as a glass slide, which is coated with an aldehyde or another reactive group which will form a covalent link with the reactive group that is on the amplified product and become covalently attached to the glass slide. Microarrays comprising the amplified products can be fabricated using a Biodot (BioDot, Inc. Irvine, CA) spotting apparatus and aldehyde-coated glass slides (CEL Associates, Houston, TX). Amplification products can be spotted onto the aldehyde-coated slides, and processed according to published procedures (Schena et al., *Proc. Natl. Acad. Sci. U.S.A.* (1995) 93:10614-10619). Arrays can also be printed by robotics onto glass, nylon (Ramsay, G., *Nature Biotechnol.* (1998), 16:40-44), polypropylene (Matson, et al., *Anal Biochem.* (1995), 224(1):110-6), and silicone slides (Marshall, A. and Hodgson, J., *Nature Biotechnol.* (1998), 16:27-31). Other approaches to array assembly include fine micropipetting within electric fields (Marshall and Hodgson, *supra*), and spotting the polynucleotides directly onto positively coated plates. Methods such as those using amino propyl silicon surface chemistry are also known in the art, as disclosed at [www.cmt.corning.com](http://www.cmt.corning.com) and <http://cmgm.stanford.edu/pbrown/>.

**[0063]** One method for making microarrays is by making high-density polynucleotide arrays. Techniques are known for rapid deposition of polynucleotides (Blanchard et al., *Biosensors & Bioelectronics*, 11:687-690). Other methods for making microarrays, e.g., by masking (Maskos and Southern, *Nuc. Acids. Res.* (1992), 20:1679-1684), may also be used. In principle, and as noted above, any type of array, for example, dot blots on a nylon hybridization membrane, could be used. However, as will be recognized by those skilled in the art, very small arrays will frequently be preferred because hybridization volumes will be smaller.

**[0064]** In one embodiment, the invention provides an array comprising at least two isolated polynucleotide molecules, wherein each isolated polynucleotide molecule detects an expressed gene product of a gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927, and wherein the gene is differentially expressed in atherosclerotic disease in a mammal. In one embodiment, the invention provides an array comprising at least two isolated polynucleotide molecules, wherein each isolated polynucleotide molecule detects an expressed gene product of a gene selected from the group of

genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs:1-927, and wherein the gene is differentially expressed in atherosclerotic disease in a mammal. In various embodiments, an array in accordance with the invention comprises any of at least 2, 3, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, or 100 polynucleotides each comprising at least a portion of a polynucleotide depicted in the Sequence Listing or a polynucleotide complement thereof.

**[0065]** In another embodiment, the invention provides an array comprising at least two antibody molecules or antigen binding fragments thereof, wherein each antibody molecule or antigen binding fragment thereof detects an expressed gene product of a gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927, and wherein the gene is differentially expressed in atherosclerotic disease in a mammal. In another embodiment, the invention provides an array comprising at least two antibody molecules or antigen binding fragments thereof, wherein each antibody molecule or antigen binding fragment thereof detects an expressed gene product of a gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs:1-927, and wherein the gene is differentially expressed in atherosclerotic disease in a mammal. In various embodiments, an antibody array in accordance with the invention comprises any of at least 2, 3, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, or 100 antibodies or antigen binding fragments thereof each recognizing an expression product (*e.g.*, a polypeptide) of a gene corresponding to a polynucleotide sequence depicted in the Sequence Listing.

### ***Methods of the invention***

#### ***Methods for detecting gene expression***

**[0066]** The invention provides methods for detecting gene expression, comprising contacting products of gene expression (*e.g.*, mRNA, protein) in a sample with a system for detecting gene expression as described above, and detecting interaction between the products of gene expression in the sample and the system for detecting gene expression. The methods for detecting gene expression described herein may be used to detect or quantify differential expression and/or for expression profiling of a sample. As used herein, “differential expression” refers to increased (upregulated) or decreased (downregulated) production of an

expressed product of a gene (e.g., mRNA, protein). Differential expression may be assessed qualitatively (presence or absence of a gene product) and/or quantitatively (change in relative amount, *i.e.*, increase or decrease, of a gene product).

**[0067]** In one embodiment, mRNA from a sample is contacted with a system for detecting gene expression comprising isolated polynucleotide molecules as described above, and hybridization complexes formed, if any, between the mRNA in the sample and the polynucleotide sequences of the system for detecting gene expression, are detected. In other embodiments, the mRNA is converted to nucleic acid derived from the mRNA, for example, cDNA, and/or amplified, prior to contact with the system for detecting gene expression.

**[0068]** In another embodiment, polypeptides from a sample are contacted with a system for detecting gene expression comprising antibodies or antigen fragments thereof that bind to polypeptide expression products of genes corresponding to the polynucleotide sequences described herein, and binding between the antibodies and polypeptides in the sample, if any, is detected.

#### *Methods for expression profiling*

**[0069]** An “expression profile” or “molecular signature” is a representation of gene expression in a sample, for example, evaluation of presence, absence, or amount of a plurality of gene expression products, such as mRNA transcripts, or polypeptide translation products of mRNA transcripts. Expression patterns constitute a set of relative or absolute expression values for a number of RNA or protein products corresponding to the plurality of genes evaluated, referred to as the subject’s “expression profile” for those nucleotide sequences. In various embodiments, expression patterns corresponding to at least about 2, 5, 10, 20, 30, 50, 100, 200, or 500, or more nucleotide sequences are obtained. The expression pattern for each differentially expressed component member of the expression profile may provide a specificity and sensitivity with respect to predictive value, *e.g.*, for diagnosis, prognosis, monitoring treatment, etc. In some embodiments, a molecular signature is determined by a statistical algorithm that determines the optimal relation between patterns of expression for various genes.

**[0070]** In some embodiments, an expression profile from an individual is compared with a reference expression profile to determine, for example, presence or absence of a disease

condition, symptom, or criterion, extent of progression of disease, effectiveness of treatment of disease, or prognosis for prophylaxis, therapy, or cure of disease.

[0071] As used herein, the term “subject” refers to an individual regardless of health and/or disease status. For example, a subject may be a patient, a study participant, a control subject, a screening subject, or any other class of individual from whom a sample is obtained and assessed in the context of the invention. Accordingly, a subject may be diagnosed with a disease, can present with one or more symptom of a disease, or may have a predisposing factor, such as a genetic or medical history factor, for a disease. Alternatively, a subject may be healthy with respect to any of the aforementioned disease factors or criteria. It will be appreciated that the term “healthy” as used herein, is relative to a specified disease condition, factor, or criterion. Thus, an individual described as healthy with reference to any specified disease or disease criterion, can be diagnosed with any other one or more disease, or may exhibit any other one or more disease criterion.

#### *Methods for Obtaining Expression Data*

[0072] Numerous methods for obtaining expression data are known, and any one or more of these techniques, singly or in combination, are suitable for determining expression profiles in the context of the present invention. For example, expression patterns can be evaluated by northern analysis, PCR, RT-PCR, Taq Man analysis, FRET detection, monitoring one or more molecular beacon, hybridization to an oligonucleotide array, hybridization to a cDNA array, hybridization to a polynucleotide array, hybridization to a liquid microarray, hybridization to a microelectric array, molecular beacons, cDNA sequencing, clone hybridization, cDNA fragment fingerprinting, serial analysis of gene expression (SAGE), subtractive hybridization, differential display and/or differential screening (see, *e.g.*, Lockhart and Winzeler (2000) *Nature* 405:827-836, and references cited therein).

[0073] For example, specific PCR primers are designed to a member(s) of a candidate nucleotide library (*e.g.*, a polynucleotide member of a system for detecting gene expression). cDNA is prepared from subject sample RNA by reverse transcription from a poly-dT oligonucleotide primer, and subjected to PCR. Double stranded cDNA may be prepared using primers suitable for reverse transcription of the PCR product, followed by amplification of the cDNA using *in vitro* transcription. The product of *in vitro* transcription is a sense-RNA

corresponding to the original member(s) of the candidate library. PCR product may be also be evaluated in a number of ways known in the art, including real-time assessment using detection of labeled primers, *e.g.* TaqMan or molecular beacon probes. Technology platforms suitable for analysis of PCR products include the ABI 7700, 5700, or 7000 Sequence Detection Systems (Applied Biosystems, Foster City, Calif.), the MJ Research Opticon (MJ Research, Waltham, Mass.), the Roche Light Cycler (Roche Diagnostics, Indianapolis, Ind.), the Stratagene MX4000 (Stratagene, La Jolla, Calif.), and the Bio-Rad iCycler (Bio-Rad Laboratories, Hercules, Calif.). Alternatively, molecular beacons are used to detect presence of a nucleic acid sequence in an unamplified RNA or cDNA sample, or following amplification of the sequence using any method, *e.g.*, IVT (in vitro transcription) or NASBA (nucleic acid sequence based amplification). Molecular beacons are designed with sequences complementary to member(s) of a candidate nucleotide library, and are linked to fluorescent labels. Each probe has a different fluorescent label with non-overlapping emission wavelengths. For example, expression of ten genes may be assessed using ten different sequence-specific molecular beacons.

**[0074]** Alternatively, or in addition, molecular beacons are used to assess expression of multiple nucleotide sequences simultaneously. Molecular beacons with sequences complimentary to the members of a diagnostic nucleotide set are designed and linked to fluorescent labels. Each fluorescent label used must have a non-overlapping emission wavelength. For example, 10 nucleotide sequences can be assessed by hybridizing 10 sequence specific molecular beacons (each labeled with a different fluorescent molecule) to an amplified or non-amplified RNA or cDNA sample. Such an assay bypasses the need for sample labeling procedures.

**[0075]** Alternatively, or in addition, bead arrays can be used to assess expression of multiple sequences simultaneously (see, *e.g.*, LabMAP 100, Luminex Corp, Austin, Tex.). Alternatively, or in addition, electric arrays can be used to assess expression of multiple sequences, as exemplified by the e-Sensor technology of Motorola (Chicago, Ill.) or Nanochip technology of Nanogen (San Diego, Calif.).

**[0076]** Of course, the particular method elected will be dependent on such factors as quantity of RNA recovered, practitioner preference, available reagents and equipment, detectors, and the like. Typically, however, the elected method(s) will be appropriate for

processing the number of samples and probes of interest. Methods for high-throughput expression analysis are discussed below.

[0077] Alternatively, expression at the level of protein products of gene expression is performed. For example, protein expression in a sample can be evaluated by one or more method selected from among: western analysis, two-dimensional gel analysis, chromatographic separation, mass spectrometric detection, protein-fusion reporter constructs, colorimetric assays, binding to a protein array (*e.g.*, antibody array), and characterization of polysomal mRNA. One particularly favorable approach involves binding of labeled protein expression products to an array of antibodies specific for members of the candidate library. Methods for producing and evaluating antibodies are well known in the art, see, *e.g.*, Coligan, *supra*; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY (“Harlow and Lane”). Additional details regarding a variety of immunological and immunoassay procedures adaptable to the present invention by selection of antibody reagents specific for the products of candidate nucleotide sequences can be found in, *e.g.*, Stites and Terr (eds.) (1991) *Basic and Clinical Immunology*, 7th ed. Another approach uses systems for performing desorption spectrometry. Commercially available systems, *e.g.*, from Ciphergen Biosystems, Inc. (Fremont, Calif.) are particularly well suited to quantitative analysis of protein expression. Protein Chip.RTM. arrays (see, *e.g.*, the website, [ciphergen.com](http://ciphergen.com)) used in desorption spectrometry approaches provide arrays for detection of protein expression. Alternatively, affinity reagents, (*e.g.*, antibodies, small molecules, etc.) may be developed that recognize epitopes of one or more protein products. Affinity assays are used in protein array assays, *e.g.*, to detect the presence or absence of particular proteins. Alternatively, affinity reagents are used to detect expression using the methods described above. In the case of a protein that is expressed on a cell surface, labeled affinity reagents are bound to a sample, and cells expressing the protein are identified and counted using fluorescent activated cell sorting (FACS).

#### *High Throughput Expression Assays*

[0078] A number of suitable high throughput formats exist for evaluating gene expression. Typically, the term high throughput refers to a format that performs at least about 100 assays, or at least about 500 assays, or at least about 1000 assays, or at least about 5000 assays, or at least

about 10,000 assays, or more per day. When enumerating assays, either the number of samples or the number of candidate nucleotide sequences evaluated can be considered. For example, a northern analysis of, *e.g.*, about 100 samples performed in a gridded array, *e.g.*, a dot blot, using a single probe corresponding to a polynucleotide sequence as described herein can be considered a high throughput assay. More typically, however, such an assay is performed as a series of duplicate blots, each evaluated with a distinct probe corresponding to a different polynucleotide sequence of a system for detecting gene expression. Alternatively, methods that simultaneously evaluate expression of about 100 or more polynucleotide sequences in one or more samples, or in multiple samples, are considered high throughput.

**[0079]** Numerous technological platforms for performing high throughput expression analysis are known. Generally, such methods involve a logical or physical array of either the subject samples, or the candidate library, or both. Common array formats include both liquid and solid phase arrays. For example, assays employing liquid phase arrays, *e.g.*, for hybridization of nucleic acids, binding of antibodies or other receptors to ligand, etc., can be performed in multiwell, or microtiter, plates. Microtiter plates with 96, 384 or 1536 wells are widely available, and even higher numbers of wells, *e.g.*, 3456 and 9600 can be used. In general, the choice of microtiter plates is determined by the methods and equipment, *e.g.*, robotic handling and loading systems, used for sample preparation and analysis. Exemplary systems include, *e.g.*, the ORCA.TM. system from Beckman-Coulter, Inc. (Fullerton, Calif.) and the Zymate systems from Zymark Corporation (Hopkinton, Mass.).

**[0080]** Alternatively, a variety of solid phase arrays can favorably be employed to determine expression patterns in the context of the invention. Exemplary formats include membrane or filter arrays (*e.g.*, nitrocellulose, nylon), pin arrays, and bead arrays (*e.g.*, in a liquid "slurry"). Typically, probes corresponding to nucleic acid or protein reagents that specifically interact with (*e.g.*, hybridize to or bind to) an expression product corresponding to a member of the candidate library, are immobilized, for example by direct or indirect cross-linking, to the solid support. Essentially any solid support capable of withstanding the reagents and conditions necessary for performing the particular expression assay can be utilized. For example, functionalized glass, silicon, silicon dioxide, modified silicon, any of a variety of polymers, such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, polycarbonate, or combinations thereof can all serve as the substrate for a solid phase array.

**[0081]** In one embodiment, the array is a “chip” composed, *e.g.*, of one of the above-specified materials. Polynucleotide probes, *e.g.*, RNA or DNA, such as cDNA, synthetic oligonucleotides, and the like, or binding proteins such as antibodies or antigen-binding fragments or derivatives thereof, that specifically interact with expression products of individual components of the candidate library are affixed to the chip in a logically ordered manner, *i.e.*, in an array. In addition, any molecule with a specific affinity for either the sense or anti-sense sequence of the marker nucleotide sequence (depending on the design of the sample labeling), can be fixed to the array surface without loss of specific affinity for the marker and can be obtained and produced for array production, for example, proteins that specifically recognize the specific nucleic acid sequence of the marker, ribozymes, peptide nucleic acids (PNA), or other chemicals or molecules with specific affinity.

**[0082]** Detailed discussion of methods for linking nucleic acids and proteins to a chip substrate, are found in, *e.g.*, U.S. Pat. No. 5,143,854, “Large Scale Photolithographic Solid Phase Synthesis Of Polypeptides And Receptor Binding Screening Thereof,” to Pirrung et al., issued, Sep. 1, 1992; U.S. Pat. No. 5,837,832, “Arrays Of Nucleic Acid Probes On Biological Chips,” to Chee et al., issued Nov. 17, 1998; U.S. Pat. No. 6,087,112, “Arrays With Modified Oligonucleotide And Polynucleotide Compositions,” to Dale, issued Jul. 11, 2000; U.S. Pat. No. 5,215,882, “Method Of Immobilizing Nucleic Acid On A Solid Substrate For Use In Nucleic Acid Hybridization Assays,” to Bahl et al., issued Jun. 1, 1993; U.S. Pat. No. 5,707,807, “Molecular Indexing For Expressed Gene Analysis,” to Kato, issued Jan. 13, 1998; U.S. Pat. No. 5,807,522, “Methods For Fabricating Microarrays Of Biological Samples,” to Brown et al., issued Sep. 15, 1998; U.S. Pat. No. 5,958,342, “Jet Droplet Device,” to Gamble et al., issued Sep. 28, 1999; U.S. Pat. No. 5,994,076, “Methods Of Assaying Differential Expression,” to Chenchik et al., issued Nov. 30, 1999; U.S. Pat. No. 6,004,755, “Quantitative Microarray Hybridization Assays,” to Wang, issued Dec. 21, 1999; U.S. Pat. No. 6,048,695, “Chemically Modified Nucleic Acids And Method For Coupling Nucleic Acids To Solid Support,” to Bradley et al., issued Apr. 11, 2000; U.S. Pat. No. 6,060,240, “Methods For Measuring Relative Amounts Of Nucleic Acids In A Complex Mixture And Retrieval Of Specific Sequences Therefrom,” to Kamb et al., issued May 9, 2000; U.S. Pat. No. 6,090,556, “Method For Quantitatively Determining The Expression Of A Gene,” to Kato, issued Jul. 18,

2000; and U.S. Pat. No. 6,040,138, "Expression Monitoring By Hybridization To High Density Oligonucleotide Arrays," to Lockhart et al., issued Mar. 21, 2000.

**[0083]** For example, cDNA inserts corresponding to candidate nucleotide sequences, in a standard TA cloning vector, are amplified by a polymerase chain reaction for approximately 30-40 cycles. The amplified PCR products are then arrayed onto a glass support by any of a variety of well-known techniques, *e.g.*, the VSLIPS.TM. technology described in U.S. Pat. No. 5,143,854. RNA, or cDNA corresponding to RNA, isolated from a subject sample, is labeled, *e.g.*, with a fluorescent tag, and a solution containing the RNA (or cDNA) is incubated under conditions favorable for hybridization, with the "probe" chip. Following incubation, and washing to eliminate non-specific hybridization, the labeled nucleic acid bound to the chip is detected qualitatively or quantitatively, and the resulting expression profile for the corresponding candidate nucleotide sequences is recorded. Multiple cDNAs from a nucleotide sequence that are non-overlapping or partially overlapping may also be used.

**[0084]** In another approach, oligonucleotides corresponding to members of a candidate nucleotide library are synthesized and spotted onto an array. Alternatively, oligonucleotides are synthesized onto the array using methods known in the art, *e.g.* Hughes, et al. *supra*. The oligonucleotide is designed to be complementary to any portion of the candidate nucleotide sequence. In addition, in the context of expression analysis for, *e.g.* diagnostic use of diagnostic nucleotide sets, an oligonucleotide can be designed to exhibit particular hybridization characteristics, or to exhibit a particular specificity and/or sensitivity, as further described below.

**[0085]** Oligonucleotide probes may be designed on a contract basis by various companies (for example, Compugen, Mergen, Affymetrix, Telechem), or designed from the candidate sequences using a variety of parameters and algorithms as indicated at the website [genome.wi.mit.edu/cgi-bin/prtm-er/primer3.cgi](http://genome.wi.mit.edu/cgi-bin/prtm-er/primer3.cgi). Briefly, the length of the oligonucleotide to be synthesized is determined, preferably at least 16 nucleotides, generally 18-24 nucleotides, 24-70 nucleotides and, in some circumstances, more than 70 nucleotides. The sequence analysis algorithms and tools described above are applied to the sequences to mask repetitive elements, vector sequences and low complexity sequences. Oligonucleotides are selected that are specific to the candidate nucleotide sequence (based on a Blast n search of the oligonucleotide sequence in question against gene sequences databases, such as the Human Genome Sequence, UniGene,

dbEST or the non-redundant database at NCBI), and have <50% G content and 25-70% G+C content. Desired oligonucleotides are synthesized using well-known methods and apparatus, or ordered from a commercial supplier.

**[0086]** A hybridization signal may be amplified using methods known in the art, and as described herein, for example use of the Clontech kit (Glass Fluorescent Labeling Kit), Stratagene kit (Fairplay Microarray Labeling Kit), the Micromax kit (New England Nuclear, Inc.), the Genisphere kit (3DNA Submicro), linear amplification, *e.g.*, as described in U.S. Pat. No. 6,132,997 or described in Hughes, T R, et al. (2001) *Nature Biotechnology* 19:343-347 (2001) and/or Westin et al. (2000) *Nat Biotech.* 18:199-204. In some cases, amplification techniques do not increase signal intensity, but allow assays to be done with small amounts of RNA.

**[0087]** Alternatively, fluorescently labeled cDNA are hybridized directly to the microarray using methods known in the art. For example, labeled cDNA are generated by reverse transcription using Cy3- and Cy5-conjugated deoxynucleotides, and the reaction products purified using standard methods. It is appreciated that the methods for signal amplification of expression data useful for identifying diagnostic nucleotide sets are also useful for amplification of expression data for diagnostic purposes.

**[0088]** Microarray expression may be detected by scanning the microarray with a variety of laser or CCD-based scanners, and extracting features with numerous software packages, for example, Imagen (Biodiscovery), Feature Extraction Software (Agilent), Scanalyze (Eisen, M. 1999. SCANALYZE User Manual; Stanford Univ., Stanford, Calif. Ver 2.32.), GenePix (Axon Instruments).

**[0089]** In another approach, hybridization to microelectric arrays is performed, *e.g.*, as described in Umek et al (2001) *J Mol Diagn.* 3:74-84. An affinity probe, *e.g.*, DNA, is deposited on a metal surface. The metal surface underlying each probe is connected to a metal wire and electrical signal detection system. Unlabelled RNA or cDNA is hybridized to the array, or alternatively, RNA or cDNA sample is amplified before hybridization, *e.g.*, by PCR. Specific hybridization of sample RNA or cDNA results in generation of an electrical signal, which is transmitted to a detector. See Westin (2000) *Nat Biotech.* 18:199-204 (describing anchored multiplex amplification of a microelectronic chip array); Edman (1997) *NAR* 25:4907-14; Vignali (2000) *J Immunol Methods* 243:243-55.

*Evaluation of Expression Patterns*

**[0090]** Expression patterns can be evaluated by qualitative and/or quantitative measures. Certain of the above described techniques for evaluating gene expression (*e.g.*, as RNA or protein products) yield data that are predominantly qualitative in nature, *i.e.*, the methods detect differences in expression that classify expression into distinct modes without providing significant information regarding quantitative aspects of expression. For example, a technique can be described as a qualitative technique if it detects the presence or absence of expression of a candidate nucleotide sequence, *i.e.*, an on/off pattern of expression. Alternatively, a qualitative technique measures the presence (and/or absence) of different alleles, or variants, of a gene product.

**[0091]** In contrast, some methods provide data that characterize expression in a quantitative manner. That is, the methods relate expression on a numerical scale, *e.g.*, a scale of 0-5, a scale of 1-10, a scale of +--+++, from grade 1 to grade 5, a grade from a to z, or the like. It will be understood that the numerical, and symbolic examples provided are arbitrary, and that any graduated scale (or any symbolic representation of a graduated scale) can be employed in the context of the present invention to describe quantitative differences in nucleotide sequence expression. Typically, such methods yield information corresponding to a relative increase or decrease in expression.

**[0092]** Any method that yields either quantitative or qualitative expression data is suitable for evaluating expression of candidate nucleotide sequences in a subject sample. In some cases, *e.g.*, when multiple methods are employed to determine expression patterns for a plurality of candidate nucleotide sequences, the recovered data, *e.g.*, the expression profile, for the nucleotide sequences is a combination of quantitative and qualitative data.

**[0093]** In some embodiments, qualitative and/or quantitative expression data from a sample is compared with a reference molecular signature that is indicative of, for example, presence or absence of a disease condition, symptom, or criterion, extent of progression of disease, effectiveness of treatment of disease, or prognosis for prophylaxis, therapy, or cure of disease. The reference molecular signature may be from a reference healthy individual (*e.g.*, an individual who does not exhibit symptoms of the disease condition to be evaluated) or an individual with a disease condition for comparison with the sample (*e.g.*, an individual with the

same or different stage of disease for comparison with the individual being evaluated, or with a genotype or phenotype that indicates, for example, prognosis for successful treatment), or the reference molecular signature may be established from a compilation of data from multiple individuals

**[0094]** In some applications, expression of a plurality of candidate polynucleotide sequences is evaluated sequentially. This is typically the case for methods that can be characterized as low- to moderate throughput. In contrast, as the throughput of the elected assay increases, expression for the plurality of candidate polynucleotide sequences in a sample or multiple samples is typically assayed simultaneously. Again, the methods (and throughput) are largely determined by the individual practitioner, although, typically, it is preferable to employ methods that permit rapid, *e.g.* automated or partially automated, preparation and detection, on a scale that is time-efficient and cost-effective.

#### *Genotyping*

**[0095]** In addition to, or in conjunction with, the correlation of expression profiles and clinical data, it is often desirable to correlate expression patterns with a subject's genotype at one or more genetic loci or to correlate both expression profiles and genetic loci data with clinical data. The selected loci can be, for example, chromosomal loci corresponding to one or more member of the candidate library, polymorphic alleles for marker loci, or alternative disease related loci (not contributing to the candidate library) known to be, or putatively associated with, a disease (or disease criterion). Indeed, it will be appreciated that where a (polymorphic) allele at a locus is linked to a disease (or to a predisposition to a disease), the presence of the allele can itself be a disease criterion.

**[0096]** Numerous well known methods exist for evaluating the genotype of an individual, including southern analysis, restriction fragment length polymorphism (RFLP) analysis, polymerase chain reaction (PCR), amplification length polymorphism (AFLP) analysis, single stranded conformation polymorphism (SSCP) analysis, single nucleotide polymorphism (SNP) analysis (*e.g.*, via PCR, Taqman or molecular beacons), among many other useful methods. Many such procedures are readily adaptable to high throughput and/or automated (or semi-automated) sample preparation and analysis methods. Often, these methods can be performed on nucleic acid samples recovered via simple procedures from the same sample as yielded the

material for expression profiling. Exemplary techniques are described in, *e.g.*, Sambrook, and Ausubel, *supra*.

### *Samples*

[0097] Samples which may be evaluated for differential expression of the polynucleotide sequences described herein include any blood vessel or portion thereof with atherosclerotic and/or inflammatory disease. Such blood vessels include, but are not limited to, the aorta, a coronary artery, the carotid artery, and peripheral blood vessels such as, for example, iliac or femoral arteries. In one embodiment, the sample is derived from an arterial biopsy. In another embodiment, the sample is derived from an atherectomy. Samples may also be derived from peripheral blood cells or serum.

[0098] Samples may be stabilized for storage by addition of reagents such as Trizol. Total RNA and/or protein may be isolated using standard techniques known in the art for expression profiling experiments.

[0099] Methods for RNA isolation include those described in standard molecular biology textbooks. Commercially available kits such as those provided by Qiagen (RNeasy Kits) may also be used for RNA isolation.

### *Methods for diagnosing atherosclerotic disease*

[0100] The invention provides methods for diagnosing an atherosclerotic disease condition in an individual. Diagnosis includes, for example, determining presence or absence of a disease condition or a symptom of a disease condition in an individual who has, who is suspected of having, or who may be suspected of being predisposed to an atherosclerotic disease. In accordance with methods of the invention for diagnosing atherosclerotic disease, gene expression products (*e.g.*, RNA or proteins) from a sample from an individual are contacted with a system for detecting gene expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

[0101] In some embodiments, qualitative and/or quantitative levels of gene expression in a test sample are compared with levels of expression in a molecular signature that is indicative of presence or absence of an atherosclerotic disease condition for which diagnosis is desired. To obtain a diagnosis, the levels of gene expression in a sample may be compared to one or more than one molecular signature, each of which may be indicative of presence or absence one or more than one atherosclerotic disease condition.

[0102] In some embodiments, polynucleotides derived from a sample from an individual (e.g., mRNA or polynucleotides derived from mRNA, for example cDNA) are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes formed from at least one of the isolated polynucleotides is indicative of presence or absence of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of presence or absence of a disease condition, criterion, or symptom for which diagnosis is desired.

[0103] In some embodiments, polypeptides derived from a sample from an individual are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of presence or absence of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is compared with presence, absence, or amount of polypeptides in a molecular signature indicative of presence or absence of a disease condition, criterion, or symptom for which diagnosis is desired.

*Methods for assessing extent of progression of atherosclerotic disease*

[0104] The invention provides methods for assessing extent of progression of an atherosclerotic disease condition in an individual. For example, a stage to which a disease

condition or particular symptom has progressed may be assessed. In accordance with methods of the invention for assessing extent of progression of atherosclerotic disease, gene expression products (*e.g.*, RNA or proteins) from a sample from an individual are contacted with a system for detecting gene expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

**[0105]** In some embodiments, qualitative and/or quantitative levels of gene expression in a test sample are compared with levels of expression in a molecular signature that is indicative of extent of progression of an atherosclerotic disease condition for which assessment is desired. The levels of gene expression may be compared to one or more than one molecular signature, each of which may be indicative of extent of progression of one or more than one atherosclerotic disease condition.

**[0106]** In some embodiments, polynucleotides derived from a sample from an individual (*e.g.*, mRNA or polynucleotides derived from mRNA, for example cDNA) are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes formed from at least one of the isolated polynucleotides is indicative of extent of progression of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of extent of progression of a disease condition for which diagnosis is desired.

**[0107]** In some embodiments, polypeptides derived from a sample from an individual are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in

the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of extent of progression of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is compared with presence, absence, or amount of polypeptides in a molecular signature indicative of extent of progression of a disease condition for which diagnosis is desired.

*Methods for assessing efficacy of treatment of atherosclerotic disease*

**[0108]** The invention provides methods for assessing extent of progression of an atherosclerotic disease condition in an individual. For example, a stage to which a disease condition or particular symptom has progressed may be assessed by the methods of the invention. In accordance with methods of the invention for assessing extent of progression of atherosclerotic disease, gene expression products (*e.g.*, RNA or proteins) from a sample from an individual are contacted with the system for detecting gene expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

**[0109]** In some embodiments, qualitative and/or quantitative levels of gene expression in a test sample are compared with levels of expression in a molecular signature that is indicative of extent of progression of an atherosclerotic disease condition for which assessment is desired. The levels of gene expression may be compared to one or more than one molecular signature, each of which may be indicative of extent of progression of one or more than one atherosclerotic disease condition.

**[0110]** In some embodiments, polynucleotides derived from a sample from an individual (*e.g.*, mRNA or polynucleotides derived from mRNA, for example cDNA) are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes

formed from at least one of the isolated polynucleotides is indicative of extent of progression of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of extent of progression of a disease condition for which assessment is desired.

[0111] In some embodiments, polypeptides derived from a sample from an individual are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of extent of progression of an atherosclerotic disease in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is compared with presence, absence, or amount of polypeptides in a molecular signature indicative of extent of progression of a disease condition for which assessment is desired.

#### *Methods for assessing efficacy of treatment*

[0112] The invention provides methods for assessing efficacy of treatment of an atherosclerotic disease symptom or condition in an individual. As used herein, "efficacy of treatment" refers to achievement of a desired therapeutic outcome (e.g., reduction or elimination of one or more symptoms of atherosclerotic disease). "Treatment" as used herein may refer to prophylaxis, therapy, or cure with respect to one or more symptoms of an atherosclerotic disease or condition. Treatment includes administration of one or more compounds or biological substances with potential therapeutic benefit and/or alterations in environmental factors, such as, for example, diet and/or exercise. In one embodiment, administration of the one or more compounds or biological substances comprises administration via a medical device such as, for example, a drug eluting stent. In other embodiments, treatment may include gene therapy or any other method that alters expression of the polynucleotide sequences described herein. In accordance with methods of the invention for assessing efficacy of treatment of atherosclerotic disease, gene expression products (e.g., RNA or proteins) from a sample from an individual are contacted with a system for detecting gene

expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

**[0113]** In some embodiments, qualitative and/or quantitative levels of gene expression in a test sample are compared with levels of expression in a molecular signature that is indicative of efficacy of treatment of an atherosclerotic disease symptom or condition for which assessment is desired. The levels of gene expression may be compared to one or more than one molecular signature, each of which may be indicative of extent of effectiveness of treatment of one or more than one atherosclerotic disease symptom or condition.

**[0114]** In some embodiments, polynucleotides derived from a sample from an individual (*e.g.*, mRNA or polynucleotides derived from mRNA, for example cDNA) are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes formed from at least one of the isolated polynucleotides is indicative of efficacy of treatment of an atherosclerotic disease symptom or condition in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of efficacy of treatment of a disease symptom or condition for which assessment is desired.

**[0115]** In some embodiments, polypeptides derived from a sample from an individual are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of efficacy of treatment of an atherosclerotic disease condition in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is

compared with presence, absence, or amount of polypeptides in a molecular signature indicative of efficacy of treatment of a disease condition for which assessment is desired.

*Methods for identifying compounds effective for treatment of atherosclerotic disease*

[0116] The invention provides methods for identifying compounds effective for treatment of an atherosclerotic disease symptom or condition in an individual. In accordance with methods of the invention for identifying compounds effective for treatment of atherosclerotic disease, at least one test compound (*i.e.*, one or more than one test compound) is administered, for example as a pharmaceutical composition comprising the at least one test compound and a pharmaceutically acceptable excipient, to an individual with an atherosclerotic disease symptom or condition or suspected of having an atherosclerotic disease symptom or condition, or to an individual who is predisposed to or suspected of being predisposed to development of an atherosclerotic disease symptom or condition. Gene expression products (*e.g.*, RNA or proteins) from a sample from the individual are contacted with a system for detecting gene expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

[0117] In some embodiments, qualitative and/or quantitative levels of gene expression in a test sample from the individual to whom the at least one test compound has been administered are compared with levels of expression in a molecular signature that is indicative of efficacy of treatment of the atherosclerotic disease symptom or condition for which assessment is desired. The levels of gene expression may be compared to one or more than one molecular signature, each of which may be indicative of extent of effectiveness of treatment of one or more than one atherosclerotic disease symptom or condition.

[0118] In some embodiments, polynucleotides derived from a sample from an individual (*e.g.*, mRNA or polynucleotides derived from mRNA, for example cDNA) to whom at least one test compound has been administered are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide

molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes formed from at least one of the isolated polynucleotides is indicative of efficacy of treatment of an atherosclerotic disease symptom or condition in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of efficacy of treatment of a disease symptom or condition for which assessment is desired.

[0119] In some embodiments, polypeptides derived from a sample from an individual to whom at least one test compound has been administered are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of efficacy of treatment of an atherosclerotic disease condition in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is compared with presence, absence, or amount of polypeptides in a molecular signature indicative of efficacy of treatment of a disease condition for which assessment is desired.

#### *Methods for determining prognosis of atherosclerotic disease*

[0120] The invention provides methods for determining prognosis of atherosclerotic disease in an individual, comprising contacting polynucleotides derived from a sample from the individual with a system for detecting gene expression as described above. "Prognosis" as used herein refers to the probability that an individual will develop an atherosclerotic disease symptom or condition, or that atherosclerotic disease will progress in an individual who has an atherosclerotic disease. Prognosis is a determination or prediction of probable course and/or outcome of a disease condition, *i.e.*, whether an individual will exhibit or develop symptoms of the disease, *i.e.*, a clinical event. In cardiovascular medicine, a common measure of prognosis is (but is not limited to) MACE (major adverse cardiac event). MACE includes mortality as

well as morbidity measures, such as myocardial infarction, angina, stroke, rate of revascularization, hospitalization, etc.

**[0121]** For determination of prognosis of atherosclerotic disease, gene expression products (*e.g.*, RNA or proteins) from a sample from an individual are contacted with the system for detecting gene expression as described above. In one embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927. In another embodiment, the genes for which expression is detected are selected from the group of genes corresponding to SEQ ID NOs: 1-927.

**[0122]** In some embodiments, qualitative and/or quantitative levels of gene expression in a sample from the individual are compared with levels of expression in a molecular signature that is indicative of prognosis of the atherosclerotic disease symptom or condition for which assessment is desired. The levels of gene expression may be compared to one or more than one molecular signature, each of which may be indicative of prognosis for one or more than one atherosclerotic disease symptom or condition.

**[0123]** In some embodiments, polynucleotides derived from a sample from an individual (*e.g.*, mRNA or polynucleotides derived from mRNA, for example cDNA) are contacted with isolated polynucleotide molecules in a system for detecting gene expression as described above, wherein each isolated polynucleotide molecule detects an expressed product of a gene that is differentially expressed in atherosclerotic disease in a mammal, and hybridization complexes formed, if any, are detected, wherein presence, absence, or amount of hybridization complexes formed from at least one of the isolated polynucleotides is indicative of prognosis for development or progression an atherosclerotic disease symptom or condition in the individual. In some embodiments, presence, absence, or amount of the polynucleotides derived from the sample is compared with presence, absence, or amount of polynucleotides in a molecular signature indicative of prognosis for development or progression of a disease symptom or condition for which assessment is desired.

**[0124]** In some embodiments, polypeptides derived from a sample from an individual are contacted with a system for detecting gene expression as described above which comprises molecules capable of detectably binding to polypeptides that are differentially expressed in

atherosclerotic disease, for example, antibodies or antigen binding fragments thereof, that detect expressed polypeptide products of genes corresponding to polynucleotide sequences depicted in the Sequence Listing, wherein presence, absence, or amount of bound polypeptide is indicative of prognosis for development or progression of an atherosclerotic disease symptom or condition in the individual. In some embodiments, presence, absence, or amount of the polypeptides derived from the sample is compared with presence, absence, or amount of polypeptides in a molecular signature indicative of prognosis for development or progression of an atherosclerotic disease symptom or condition for which assessment is desired.

### ***Novel polynucleotide sequences***

**[0125]** The invention provides novel polynucleotide sequences that are differentially expressed in atherosclerotic disease. We have identified unnamed (not previously described as corresponding to a gene or an expressed gene, and/or for which no function has previously been assigned) polynucleotide sequences herein. The novel differentially expressed nucleotide sequences of the invention are useful in a system for detecting gene expression, such as a diagnostic oligonucleotide set, and are also useful as probes in a diagnostic oligonucleotide set immobilized on an array. The novel polynucleotide sequences may be useful as disease target polynucleotide sequences and/or as imaging reagents as described herein.

**[0126]** As used herein, “novel polynucleotide sequence” refers to (a) a polynucleotide sequence containing at least one of the polynucleotide sequences disclosed herein (as depicted in the Sequence Listing); (b) a polynucleotide sequence that encodes the amino acid sequence encoded by a polynucleotide sequence disclosed herein; (c) a polynucleotide sequence that hybridizes to the complement of a coding sequence disclosed herein under highly stringent conditions, *e.g.*, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65° C, and washing in 0.1x SSC/0.1% SDS at 68° C. (Ausubel, F.M. et al., eds. (1989) *Current Protocols in Molecular Biology*, Vol. 1, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, at p. 2.01.3); (d) a polynucleotide sequence that hybridizes to the complement of a coding sequence disclosed herein under less stringent conditions, such as moderately stringent conditions, *e.g.*, washing in 0.2x SSC/0.1% SDS at 42° C. (Ausubel et al. (1989), *supra*), yet which still encodes a functionally equivalent gene product; and/or (e) a polynucleotide sequence that is at least 90% identical, at least 80%

identical, or at least 70% identical to the coding sequences disclosed herein, wherein % identity is determined using standard algorithms known in the art.

[0127] The invention also includes polynucleotide molecules that hybridize to, and are therefore the complements of, novel polynucleotide molecules as described in (a) through (c) in the preceding paragraph. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the polynucleotide molecules are deoxyoligonucleotides, highly stringent conditions may refer to, *e.g.*, washing in 6x SSC/0.05% sodium pyrophosphate at 37° C (for 14-base oligonucleotides), 48° C (for 17-base oligonucleotides), 55° C (for 20-base oligonucleotides, and 60° C (for 23-base oligonucleotides). These polynucleotide molecules may act as target nucleotide sequence antisense molecules, useful, for example, in target nucleotide sequence regulation and/or as antisense primers in amplification reactions of target nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target nucleotide sequence regulation. Such molecules may also be used as components of diagnostic methods whereby the presence of a disease-causing allele may be detected.

[0128] The invention also encompasses nucleic acid molecules contained in full-length gene sequences that are related to or derived from novel polynucleotide sequences as described above and as depicted in the Sequence Listing. One sequence may map to more than one full-length gene.

[0129] The invention also encompasses (a) polynucleotide vectors that contain any of the foregoing novel polynucleotide sequences and/or their complements; (b) polynucleotide expression vectors that contain any of the foregoing novel polynucleotide sequences and/or their complements; and (c) genetically engineered host cells that contain any of the foregoing novel polynucleotide sequences operatively associated with a regulatory element that directs expression of the polynucleotide in the host cell. As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators, and other elements known to those skilled in the art that drive and regulate gene expression.

[0130] The invention includes fragments of the novel polynucleotide sequences described above. Fragments may be any of at least 5, 10, 15, 20, 25, 50, 100, 200, or 500 nucleotides, or larger.

***Novel polypeptide products***

[0131] The invention includes novel polypeptide products, encoded by genes corresponding to the novel polynucleotide sequences described above, or functionally equivalent polypeptide gene products thereof. "Functionally equivalent," as used herein, refers to a protein capable of exhibiting a substantially similar *in vivo* function, *e.g.*, activity, as a novel polypeptide gene product encoded by a novel polynucleotide of the invention.

[0132] Equivalent novel polypeptide products may include deletions, additions, and/or substitutions of amino acid residues within the amino acid sequence encoded by a gene corresponding to a novel polynucleotide sequence of the invention as described above, but which results in a "silent" change (*i.e.*, a change which does not substantially change the functional properties of the polypeptide). Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

[0133] Novel polypeptide products of genes corresponding to novel polynucleotide sequences described herein may be produced by recombinant nucleic acid technology using techniques that are well known in the art. For example, methods that are well known to those skilled in the art may be used to construct expression vectors containing novel polynucleotide coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, *supra*, and Ausubel et al., 1989, *supra*. Alternatively, RNA capable of encoding novel nucleotide sequence protein sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis" (1984) Gait, M. J. ed., IRL Press, Oxford. A variety of host-expression vector systems may be utilized to express the novel nucleotide sequence coding sequences of the invention. Ruther et al. (1983) *EMBO J.* 2:1791; Inouye & Inouye (1985) *Nucleic Acids Res.* 13:3101-3109; Van Heeke & Schuster (1989) *J. Biol. Chem.* 264:5503; Smith et al. (1983) *J. Virol.* 46: 584; Smith, U.S. Pat. No. 4,215,051; Logan & Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; Bittner et al. (1987) *Methods in Enzymol.* 153:516-544; Wigler, et al. (1977) *Cell* 11:223; Szybalska & Szybalski (1962) *Proc. Natl. Acad. Sci. USA* 48:2026; Lowy, et al. (1980) *Cell* 22:817; Wigler, et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567; O'Hare, et

al. (1981) *Proc. Natl. Acad. Sci. USA* 78:1527; Mulligan & Berg (1981) *Proc. Natl. Acad. Sci. USA* 78:2072; Colberre-Garapin, et al. (1981) *J. Mol. Biol.* 150:1; Santerre, et al. (1984) *Gene* 30:147; Janknecht, et al. (1991) *Proc. Natl. Acad. Sci. USA* 88: 8972-8976. When recombinant DNA technology is used to produce the protein encoded by a gene corresponding to the novel polynucleotide sequence, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

### ***Antibodies***

[0134] The invention also provides antibodies or antigen binding fragments thereof that specifically bind to novel polypeptide products encoded by genes that correspond to novel polynucleotide sequences as described above. Antibodies capable of specifically recognizing one or more novel nucleotide sequence epitopes may be prepared by methods that are well known in the art. Such antibodies include, but are not limited to, polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a novel polynucleotide sequence in a biological sample, or, alternatively, as a method for the inhibition of abnormal gene activity, for example, the inhibition of a disease target nucleotide sequence, as further described below. Thus, such antibodies may be utilized as part of a disease treatment method, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of novel nucleotide sequence encoded proteins, or for the presence of abnormal forms of the such proteins.

[0135] For the production of antibodies that bind to a polypeptide encoded by a novel nucleotide sequence, various host animals may be immunized by injection with a novel protein encoded by the novel nucleotide sequence, or a portion thereof. Such host animals may include, but are not limited to rabbits, mice, and rats. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

**[0136]** Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as novel polypeptide gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with novel polypeptide gene product supplemented with adjuvants as also described above.

**[0137]** Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein (1975) *Nature* 256:495-497; and U.S. Pat. No. 4,376,110, the human B-cell hybridoma technique (Kosbor et al. (1983) *Immunology Today* 4:72; and Cole et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030), and the EBV-hybridoma technique (Cole et al. (1985) *Monoclonal Antibodies And Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. A hybridoma producing a mAb may be cultivated *in vitro* or *in vivo*.

**[0138]** In addition, techniques developed for the production of "chimeric antibodies" by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. Morrison et al. (1984) *Proc. Natl. Acad. Sci.* 81:6851-6855; Neuberger et al. (1984) *Nature* 312:604-608; Takeda et al. (1985) *Nature* 314:452-454. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

**[0139]** Alternatively, techniques described for the production of single chain antibodies can be adapted to produce novel nucleotide sequence-single chain antibodies. (U.S. Pat. No. 4,946,778; Bird (1988) *Science* 242:423-426; Huston et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Ward et al. (1989) *Nature* 334:544-546) Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

**[0140]** Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')<sub>2</sub> fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments

which can be generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed (Huse et al. (1989) *Science* 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with a desired specificity.

***Disease specific target polynucleotide sequences***

[0141] The invention also provides disease specific target polynucleotide sequences, and sets of disease specific target polynucleotide sequences. The diagnostic oligonucleotide sets, individual members of the diagnostic oligonucleotide sets and subsets thereof, and novel polynucleotide sequences, as described above, may also serve as disease specific target polynucleotide sequences. In particular, individual polynucleotide sequences that are differentially regulated or have predictive value that is strongly correlated with an atherosclerotic disease or disease criterion are especially favorable as atherosclerotic disease specific target polynucleotide sequences. Sets of genes that are co-regulated may also be identified as disease specific target polynucleotide sets. Such polynucleotide sequences and/or their complements and/or the expression products of genes corresponding to such polynucleotide sequences (*e.g.*, mRNA, proteins) are targets for modulation by a variety of agents and techniques. For example, disease specific target polynucleotide sequences (or the expression products of genes corresponding to such polynucleotide sequences, or sets of disease specific target polynucleotide sequences) can be inhibited or activated by, *e.g.*, target specific monoclonal antibodies or small molecule inhibitors, or delivery of the polynucleotide sequence or an expression product of a gene corresponding to the polynucleotide sequence to patients. Also, sets of genes can be inhibited or activated by a variety of agents and techniques. The specific usefulness of the target polynucleotide sequence(s) depends on the subject groups from which they were discovered, and the disease or disease criterion with which they correlate.

***Kits***

[0142] The invention provides kits containing a system for detecting gene expression, a diagnostic nucleotide set, candidate nucleotide library, one or novel polynucleotide sequence, one or more polypeptide products of the novel polynucleotide sequences, and/or one or more antibodies that recognize polypeptide expression products of the differentially regulated

polynucleotide sequences described herein. A kit may contain a diagnostic nucleotide probe set, or other subset of a candidate library (*e.g.*, as a cDNA, oligonucleotide or antibody microarray or reagents for performing an assay on a diagnostic gene set using any expression profiling technology), packaged in a suitable container. The kit may further comprise one or more additional reagents, *e.g.*, substrates, labels, primers, reagents for labeling expression products, tubes and/or other accessories, reagents for collecting tissue or blood samples, buffers, hybridization chambers, cover slips, etc., and may also contain a software package, *e.g.*, for analyzing differential expression using statistical methods as described herein, and optionally a password and/or account number for accessing the compiled database. The kit optionally further comprises an instruction set or user manual detailing preferred methods of performing the methods of the invention, and/or a reference to a site on the Internet where such instructions may be obtained.

**Table 1. Polynucleotide sequences which detect differentially expressed genes in atherosclerotic disease**

| SEQ ID NO: | CLONE ID    | GENE SYMBOL   | GENE NAME                                                                                                                                                           | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                    | 60mer SEQUENCE                                                                                                                                                                                                                                                                                                                                                                      |
|------------|-------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.         | C0267B04-3  |               | C0267B04-5N<br>NIA Mouse<br>7.5-dpc Whole<br>Embryo cDNA<br>Library (Long)<br>Mus musculus<br>cDNA clone<br>NIA:C0267B04<br>IMAGE:30017<br>007 5', MRNA<br>sequence | C0267B04   |            | No Chromosome location<br>info available | ATGAGCC<br>ACTCACA<br>TTTTCCT<br>TCTATCA<br>AATAAGT<br>CAAGA                                                                                                                                                                                                                                                                                                                        |
| 2.         | M29697.1    | Il7r          | interleukin 7<br>receptor                                                                                                                                           | M29697     | Mm.389     | Chromosome 15                            | CCTATTG<br>GTGTCAA<br>CACCACI<br>GGATGGI<br>TAGTCCA<br>CCAAA<br>TACCTGA<br>CTCTCTA<br>TGTTGTC<br>GGCAAA<br>GCATTCC<br>TCCAAG<br>CCCTTTG<br>GTGGGC<br>TCTGAAC<br>CAAATGC<br>TAGGATC<br>CCAGA<br>TCCAAAC<br>AATGAG<br>CGCACTC<br>AGCCAT<br>ACTGAC<br>TTGGAA<br>TGCCCTTC<br>GCAACA<br>GCAGAT<br>AAGATC<br>GACACTC<br>CAGCAG<br>CATGAA<br>AACCAG<br>ATTAAC<br>CCTGAA<br>ACAATT<br>TGTGC |
| 3.         | L0304D03-3  | Wnt4          | wingless-<br>related MMTV<br>integration site<br>4                                                                                                                  | L0304D03   | Mm.103301  | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                     |
| 4.         | L0237D12-3  | Ctsd          | cathepsin D                                                                                                                                                         | L0237D12   | Mm.231395  | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                     |
| 5.         | C0266B08-3  | BM204200      | ESTs<br>BM204200                                                                                                                                                    | C0266B08   | Mm.222000  | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                     |
| 6.         | J0537C05-3  | Pfdn2         | prefoldin 2                                                                                                                                                         | J0537C05   | Mm.10756   | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                     |
| 7.         | L0216F02-3  | C430008C19Rik | RIKEN cDNA<br>C430008C19<br>gene                                                                                                                                    | L0216F02   | Mm.268474  | Chromosome 10                            |                                                                                                                                                                                                                                                                                                                                                                                     |
| 8.         | NM_017372.1 | Lyzs          | lysozyme                                                                                                                                                            | NM_017372  | Mm.45436   | Chromosome 10                            |                                                                                                                                                                                                                                                                                                                                                                                     |
| 9.         | C0271B02-3  | 4732437J24Rik | RIKEN cDNA<br>4732437J24<br>gene                                                                                                                                    | C0271B02   | Mm.39102   | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60me<br>SEQUENCE                                                                 |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------|
| 10.              | H3022C10-3 | AA408868       | expressed sequence<br>AA408868                                                                                                                                                                                             | H3022C10      | Mm.247272     | Chromosome 16            | ATTGTTG<br>CAAAG<br>CATTGGA<br>GACACGT<br>AGGCATT<br>TATTCTT<br>AGACTGT<br>TGAAT |
| 11.              | L0806E05-3 | Gtl2           | GTL2,<br>imprinted<br>maternally<br>expressed<br>untranslated<br>mRNA                                                                                                                                                      | L0806E05      | Mm.200506     | Chromosome 12            | GTAATGG<br>ATGTATC<br>CCCATAT<br>CCATCTC<br>CCTTAAC<br>TAAGCA                    |
| 12.              | H3111E06-5 | Acas2l         | acetyl-<br>Coenzyme A<br>synthetase 2<br>(AMP<br>forming)-like                                                                                                                                                             | H3111E06      | Mm.7044       | Chromosome 2             | ACACCTC<br>TCCCAAG<br>ACGAGT<br>TGTCCTC<br>TTACTTG<br>ATCATTT                    |
| 13.              | H3091H05-3 | Hrasl          | Harvey rat<br>sarcoma virus<br>oncogene 1                                                                                                                                                                                  | H3091H05      | Mm.6793       | Chromosome 7             | GTGAGAT<br>CAGCATA<br>GCGGAA<br>AACCAC<br>TGAGAGT<br>CTGGCT                      |
| 14.              | K0324B10-3 | Timp1          | tissue inhibitor<br>of<br>metalloproteinase 1                                                                                                                                                                              | K0324B10      | Mm.8245       | Chromosome X             | TCATAAG<br>AAATTCA<br>TTCCCCA<br>TCAACGA<br>ACCTTAT<br>GCGTT                     |
| 15.              | K0508B06-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>ref:NP_077285.<br>1 (H.sapiens)<br>A20-binding<br>inhibitor of NF-<br>kappaB<br>activation 2;<br>LKB1-<br>interacting<br>protein [Homo<br>sapiens] | K0508B06      | Mm.217235     | Chromosome 5             | AAAGACT<br>AGGAGTC<br>AACCAGC<br>AAACTTA<br>TGCTTTG<br>TTCCAGC                   |
| 16.              | C0176A01-3 | Syngr1         | synaptogyrin 1                                                                                                                                                                                                             | C0176A01      | Mm.230301     | Chromosome 15            | GCAGCA<br>TCCTTGC<br>TTCTTTG<br>GTTCTT<br>AAACAT<br>GAGC                         |
| 17.              | J0748G02-3 |                | AU018093<br>Mouse two-cell<br>stage embryo<br>cDNA Mus<br>musculus<br>cDNA clone<br>J0748G02 3',<br>MRNA<br>sequence                                                                                                       | J0748G02      |               | Chromosome 2             | TTTAAAC<br>CCTGAA<br>CAGGTT<br>TTTAAAC<br>ATAAAA<br>AAATAA                       |
| 18.              | J0035G10-3 | C77672         | ESTs C77672                                                                                                                                                                                                                | J0035G10      | Mm.36571      | Chromosome 4             | TAGCAT<br>ACCATG<br>CAATAC<br>TTTAGA                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                     | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------|------------|----------------|--------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 19.              | C0630C02-3 | Cxcl16         | chemokine (C-X-C motif)<br>ligand 16             | C0630C02      | Mm.46424      | Chromosome 11                            | TTAATGG<br>GAGAG<br>CCTGAGC<br>TGTTTCT<br>CTGTCTT<br>CAAAGTA<br>TATGGAA<br>GGTTA<br>GCTGGTC<br>TGTC AAC<br>ATGGCTC<br>TTGTTTC<br>CTGTAGC<br>TTGAAC<br>ACTTAAC<br>CTGCATA<br>CAATCCT<br>GGTTTGC<br>TGTCTCG<br>TCTAA<br>GGGCAA<br>ACTTTCT<br>AACTGAC<br>TGAGATC<br>CCCAAG<br>GAAAAA<br>ACTATGC<br>GGACAG<br>ATTACCA<br>GACTAA<br>ATATTCT<br>CTTTGGC<br>TCACTGA<br>AACCCCT<br>CAGAGA<br>TGAAGA<br>AAAGCT<br>GTCCAA<br>GATATA<br>ATAAAG<br>AAAGGA<br>CTGGCTC<br>AGATAC<br>GGAAC<br>CTGAAC<br>TTAATA<br>GGATAT<br>TCTTCA<br>GGATAG<br>TGAAG<br>TTTTGTT<br>CATCTGT<br>CGTTCTC<br>CTGAAT<br>TTGTCA<br>AAAA<br>GCCAAT<br>AACCCA<br>AAGGGT<br>GTATTA<br>GTTTCA<br>GCACA<br>AGTGCA<br>TGGTTTC<br>TGTGCT<br>GGTTTA |
| 20.              | K0313A10-3 | 5430435G22Rik  | RIKEN cDNA<br>5430435G22<br>gene                 | K0313A10      | Mm.44508      | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 21.              | L0070E11-3 | Cbfa2t1h       | CBFA2T1<br>identified gene<br>homolog<br>(human) | L0070E11      | Mm.4909       | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 22.              | H3072E02-3 | BG069076       | ESTs<br>BG069076                                 | H3072E02      | Mm.26437      | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 23.              | H3079B06-3 |                | Mus musculus<br>unknown<br>mRNA                  | H3079B06      | Mm.295683     | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 24.              | H3002D08-3 | 4833412N02Rik  | RIKEN cDNA<br>4833412N02<br>gene                 | H3002D08      | Mm.195099     | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 25.              | H3159A08-3 | Gp49b          | glycoprotein 49<br>B                             | H3159A08      | Mm.196617     | Chromosome 10                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 26.              | C0612F12-3 | BM207436       | ESTs<br>BM207436                                 | C0612F12      | Mm.260421     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 27.              | H3108A03-3 | Apobec1        | apolipoprotein<br>B editing<br>complex 1         | H3108A03      | Mm.3333       | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 28.              | C0180G01-3 | BI076556       | ESTs BI076556                                    | C0180G01      | Mm.37657      | Chromosome 16                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 29.              | C0938A03-3 | Sf3a1          | splicing factor<br>3a, subunit 1                 | C0938A03      | Mm.156914     | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60mer<br>SEQUENCE                                                                                                           |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| 30.              | J0703E02-3 | Ogdh           | oxoglutarate<br>dehydrogenase<br>(lipoamide)                                                                                                                       | J0703E02      | Mm.30074      | Chromosome 11                   | CTGAAAC<br>CACACT<br>CATGAGT<br>TGTGAAC<br>GACCCAC<br>TGATACT<br>TTCTGCA<br>GGGCA                                           |
| 31.              | C0274D12-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | C0274D12      | Mm.217705     | Chromosome 12                   | TAGACG1<br>AAAAGG1<br>AAGTTTA<br>TTTGTT1<br>AATCCG1<br>TGTGGG                                                               |
| 32.              | H3097H03-3 | Expi           | extracellular<br>proteinase<br>inhibitor                                                                                                                           | H3097H03      | Mm.1650       | Chromosome 11                   | ACTGTGC<br>AGCTTCC<br>GTGTTTG<br>TAAAAAT<br>ATCCTTA<br>CCTTC                                                                |
| 33.              | H3074D10-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus]                    | H3074D10      | Mm.103987     | Chromosome 15                   | TATAAA1<br>AGTGAA1<br>AACCTAC<br>GTATCTA<br>AACACT1<br>TTCAG                                                                |
| 34.              | M14222.1   | Ctsb           | cathepsin B                                                                                                                                                        | M14222        | Mm.22753      | Chromosome 14                   | CATCCTA<br>GAGGAT<br>ACTTTGC<br>ACTTCCI<br>CGTGTC1<br>GTGTGA<br>CCTGAA1<br>GTCATG1<br>CTTGGA1<br>AGTAAC<br>ACAGCT1<br>CTAGT |
| 35.              | C0176G01-3 | 2400006H24Rik  | RIKEN cDNA<br>2400006H24<br>gene                                                                                                                                   | C0176G01      | Mm.143774     | Chromosome Multiple<br>Mappings | AGTCAA1<br>CCTAAA1<br>TTATGTC<br>AGACCA<br>AGATAC<br>TGAGCA                                                                 |
| 36.              | H3092F08-5 |                | UNKNOWN:<br>Similar to Mus<br>musculus<br>immediate-<br>early antigen<br>(E-beta) gene,<br>partial intron 2<br>sequence                                            | H3092F08      |               | Chromosome 17                   | TTATGC1<br>TTTCAC1<br>AAAGGG<br>GGAGCC<br>TTGTCC1<br>TTGTAG<br>GTAACC<br>GCCCTG<br>GGAATT<br>TAGTAG<br>GGGAAA               |
| 37.              | H3054F02-3 | 1200003C15Rik  | RIKEN cDNA<br>1200003C15<br>gene                                                                                                                                   | H3054F02      | Mm.19325      | Chromosome 10                   |                                                                                                                             |
| 38.              | C0012F07-3 | 3010021M21Rik  | RIKEN cDNA<br>3010021M21<br>gene                                                                                                                                   | C0012F07      | Mm.100525     | Chromosome 9                    |                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                           | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                                                                                                                                                                 |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|----------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 39.              | L0955A10-3 | 9030409G11Rik  | RIKEN cDNA<br>9030409G11<br>gene                                                                                                                       | L0955A10      | Mm.32810       | Chromosome 4                             | TGCTCTT<br>TCCCATG<br>CCCAGAC<br>ATTTTAA<br>GTAACA/<br>ATGCTTG<br>TGAAGCT<br>AGGACA/<br>CCAGATC<br>AAGAAAG/<br>GAGCCTC<br>CACCTCC<br>CCTAAA                                                                       |
| 40.              | L0045B05-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus] | L0045B05      | Mm.182645      | Chromosome 9                             |                                                                                                                                                                                                                   |
| 41.              | H3049A10-3 | BG066966       | ESTs<br>BG066966                                                                                                                                       | H3049A10      | Mm.262549      | Chromosome 6                             | TCCTGTG<br>ATCCCAT<br>CCTGAAC<br>CGTAGTC<br>TTTCCA<br>ATTCT<br>CGACGAC<br>TTCCAAC<br>CCTGCTC<br>TGAACCC<br>TCAAAC/<br>GAGCAT<br>GAAGAG.<br>AAGATG/<br>GCCTTG/<br>AGCCAC/<br>GCAAAG/<br>AGAACA/<br>GAGGCT          |
| 42.              | X70298.1   | Sox4           | SRY-box<br>containing gene<br>4                                                                                                                        | X70298        | Mm.253853      | Chromosome 13                            |                                                                                                                                                                                                                   |
| 43.              | L0001C09-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus]        | L0001C09      | Mm.171544      | Chromosome 12                            |                                                                                                                                                                                                                   |
| 44.              | H3010D12-5 |                | UNKNOWN:<br>Similar to Mus<br>musculus<br>RIKEN cDNA<br>8430421I07<br>gene<br>(8430421I07Ri<br>k), mRNA                                                | H3010D12      | Data not found | Chromosome 9                             | GCCTGCA/<br>GTTTGTG<br>TAGCCTC<br>GAGCTG/<br>GTGCTG/<br>CCAGGC                                                                                                                                                    |
| 45.              | C0923E12-3 | Ptpns1         | protein tyrosine<br>phosphatase,<br>non-receptor<br>type substrate 1                                                                                   | C0923E12      | Mm.1682        | Chromosome 2                             | CTGTCTT<br>TTCCAA/<br>TGGTTGC<br>GCTCCAC<br>TTTTCT<br>CTAAA<br>TTCACAC<br>CCTGGTC<br>ATGCAG.<br>GAACAA<br>CTCAGG/<br>CTGGAA<br>TCTACA/<br>GCATTG/<br>CCAAGA/<br>CTTGGA/<br>TTCACTC<br>TTCITT<br>GGGCCT/<br>ATGGCT |
| 46.              | C0941E09-3 | D330001F17Rik  | RIKEN cDNA<br>D330001F17<br>gene                                                                                                                       | C0941E09      | Mm.123240      | No Chromosome location<br>info available |                                                                                                                                                                                                                   |
| 47.              | K0534C04-3 | Tce1           | T-complex<br>expressed gene<br>1                                                                                                                       | K0534C04      | Mm.41932       | Chromosome 17                            |                                                                                                                                                                                                                   |
| 48.              | H3064E11-3 | BG068354       | ESTs<br>BG068354                                                                                                                                       | H3064E11      | Mm.173544      | Chromosome 4                             |                                                                                                                                                                                                                   |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                       |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|---------------------------------------------------------------------------------------------------------|
| 49.              | L0957C02-3 | E130319B15Rik  | RIKEN cDNA<br>E130319B15<br>gene                                                                   | L0957C02      | Mm.149539     | Chromosome 2             | TACATTA<br>GTTAACA<br>TCACACA<br>GGAGC<br>TGTGTTG<br>TTCAACT<br>AGACGCC<br>ATGTCCA<br>GGAAAA/<br>AATAAA |
| 50.              | L0240C12-3 | C1qa           | complement<br>component 1, q<br>subcomponent,<br>alpha<br>polypeptide                              | L0240C12      | Mm.370        | Chromosome 4             | ACTGATG<br>TGCACAC<br>CAGTGGT<br>TTAAGCA<br>CTGGAA1<br>GATCC                                            |
| 51.              | J0018H07-3 | Rnf149         | ring finger<br>protein 149                                                                         | J0018H07      | Mm.28614      | Chromosome 1             | TCACAGA<br>TGTGGAC<br>TGTTTTC<br>TACTAGA<br>CCTCTGT<br>ATAAA                                            |
| 52.              | K0508E12-3 | Rin3           | Ras and Rab<br>interactor 3                                                                        | K0508E12      | Mm.24145      | Chromosome 12            | TCGGGG/<br>AGCTGAC<br>TCCACCA<br>CCAAAGA/<br>GAGTATT<br>TGAAGA                                          |
| 53.              | L0208A01-3 | 4933437K13Rik  | RIKEN cDNA<br>4933437K13<br>gene                                                                   | L0208A01      | Mm.159218     | Chromosome 16            | GGAGAC<br>GCTTTTA<br>TTAATGT<br>GATATTC<br>ACAAGG<br>AATGGTC                                            |
| 54.              | C0239G03-3 | BM202478       | EST<br>BM202478                                                                                    | C0239G03      | Mm.217664     | Chromosome 2             | AACTGTC<br>TAATTGT<br>GCCTGA/<br>CCAGAAC<br>AGAAAC<br>CTGGGA                                            |
| 55.              | L0518C11-3 | 1700016K05Rik  | RIKEN cDNA<br>1700016K05<br>gene                                                                   | L0518C11      | Mm.221743     | Chromosome 17            | GTGTTGT<br>GTCGTCC<br>TTAATGA<br>ACCTGAC<br>CAGTTAC<br>TTACCC                                           |
| 56.              | H3054C09-3 | Oas1c          | 2'-5'<br>oligoadenylate<br>synthetase 1C                                                           | H3054C09      | Mm.206775     | Chromosome 5             | CTATATC<br>GAGAAA/<br>ACGTATC<br>ACCCCA/<br>ACAACA<br>CTACGCC                                           |
| 57.              | L0811E07-3 | 3110057O12Rik  | RIKEN cDNA<br>3110057O12<br>gene                                                                   | L0811E07      | Mm.32373      | Chromosome 3             | GGAATA<br>TGTAGA<br>CTGGCC<br>CTTGTGC<br>CTGATGC<br>CCTCC                                               |
| 58.              | J0948A06-3 |                | Mus musculus<br>mRNA similar<br>to RIKEN<br>cDNA<br>4930503E14<br>gene (cDNA<br>clone<br>MGC:58418 | J0948A06      | Mm.261771     | Chromosome 14            | TTGGGT<br>ATATTTI<br>ACCCAT<br>CAAAAG<br>CCTACTI<br>TTCTCT                                              |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                        | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]              | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                   |
|------------------|-------------|----------------|-------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|---------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 59.              | C0931B05-3  |                | IMAGE:670814), complete cds transcribed sequence with weak similarity to protein ref:NP_081764.1 (M.musculus) RIKEN cDNA 5730493B19 | C0931B05      | Mm.252843     | Chromosome 10                         | GTTCTCTG/<br>TCTTGAT/<br>TAGGACA<br>CCCACCA<br>AAAAATGA<br>GGAATTT                                                                                                                                                                                                                                                                 |
| 60.              | H3022A09-3  | Eps8l2         | [Mus musculus] EPS8-like 2                                                                                                          | H3022A09      | Mm.27451      | Chromosome 7                          | TGACTTC.<br>GTCCCAT<br>CCCAAAG<br>CTGTGAT<br>GATGTCT<br>CTATAT<br>TGGGTAG<br>CTAGGTC<br>TGATATC<br>CTACAGT<br>CTGTAGC<br>TGACA<br>CCTGTCT<br>ACTCAAA<br>AAATCCA<br>ATCTTCA<br>CACTTTG<br>CCTAC<br>TACTCCC<br>GACTAG/<br>GTGGCTA<br>GGAGCA/<br>CAGAGC/<br>GACTGAT<br>GGGACA/<br>AAGTCA/<br>GACCACC<br>ACCAGCC<br>GAGATC/<br>TGACCA/ |
| 61.              | G0118B03-3  | Usf2           | upstream transcription factor 2                                                                                                     | G0118B03      | Mm.15781      | Chromosome 7                          |                                                                                                                                                                                                                                                                                                                                    |
| 62.              | H3156C12-3  | Ms4a6d         | membrane-spanning 4-domains, subfamily A, member 6D                                                                                 | H3156C12      | Mm.170657     | Chromosome 19                         |                                                                                                                                                                                                                                                                                                                                    |
| 63.              | H3074G06-3  | 9530020G05Rik  | RIKEN cDNA 9530020G05 gene                                                                                                          | H3074G06      | Mm.152120     | Chromosome 6                          |                                                                                                                                                                                                                                                                                                                                    |
| 64.              | NM_003254.1 | TIMP1          | tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)                                    | NM_003254     | Hs.5831       | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                    |
| 65.              | K0647H07-3  | Il7r           | interleukin 7 receptor                                                                                                              | K0647H07      | Mm.389        | Chromosome 15                         | GAAAAC/<br>ACTCTTG<br>GAGACA/<br>GCAAAA/<br>GATGTCA/<br>CTATGTC<br>TCAAGG/<br>GTAGAC/<br>GGCAGA/<br>CGTAAC/<br>GGCTCA/<br>CATCCTC<br>CACCAC/<br>ACAACC/<br>GCCATG/<br>TTTCCG/<br>CTTCTG/<br>AAGCA<br>GTACCC/<br>CTGTAT/<br>AATCGG/<br>CCTGAT/                                                                                      |
| 66.              | J0257F12-3  | Rnf25          | ring finger protein 25                                                                                                              | J0257F12      | Mm.86910      | Chromosome 1                          |                                                                                                                                                                                                                                                                                                                                    |
| 67.              | H3083G02-3  | Lcn2           | lipocalin 2                                                                                                                         | H3083G02      | Mm.9537       | Chromosome 2                          |                                                                                                                                                                                                                                                                                                                                    |
| 68.              | M64086.1    | Serpina3n      | serine (or cysteine) proteinase inhibitor, clade                                                                                    | M64086        | Mm.22650      | Chromosome 12                         |                                                                                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                                     |
|------------------|------------|----------------|---------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|--------------------------------------------------------------------------------------|
|                  |            |                | A, member 3N                                                                    |               |               |                                          | TCTTTGAC                                                                             |
| 69.              | C0906B05-3 | Cenpc          | centromere<br>autoantigen C                                                     | C0906B05      | Mm.221600     | Chromosome 5                             | GAAAC<br>AAGAACT<br>ATACAGA<br>ACTTCAG<br>TCAGTTA/<br>CTTTTAA/<br>TCTCTC             |
| 70.              | H3094B08-3 | BG071051       | ESTs<br>BG071051                                                                | H3094B08      | Mm.173358     | Chromosome 2                             | CTTGACC<br>GATGGAA<br>TACCTAG<br>GAGAAGC<br>CAAACATA<br>CTGTCA                       |
| 71.              | K0110F02-3 | Pstpip1        | proline-serine-<br>threonine<br>phosphatase-<br>interacting<br>protein 1        | K0110F02      | Mm.2534       | Chromosome 9                             | GGAACGC<br>ACGTGGC<br>TCCCTGG<br>TACTTGG<br>GCTCTGA<br>AGGCTA                        |
| 72.              | L0072G08-3 | Renbp          | renin binding<br>protein                                                        | L0072G08      | Mm.28280      | Chromosome X                             | TTCGAAT<br>ATCATTG<br>GTTTCTC<br>TGCCTTT<br>TCTGGAT<br>CCCTG                         |
| 73.              | J0088G06-3 | 4930472G13Rik  | RIKEN cDNA<br>4930472G13<br>gene                                                | J0088G06      | Mm.23172      | No Chromosome location<br>info available | GCCTGGA<br>GAAGGC/<br>TACAAAC<br>AACTTAC<br>CTATTCA<br>CTTTTG                        |
| 74.              | K0121F05-3 | Fcgr2b         | Fc receptor,<br>IgG, low<br>affinity IIb                                        | K0121F05      | Mm.10809      | Chromosome 1                             | CTGGATC<br>AAACAG/<br>TGATTAC<br>ACCACA/<br>TCTCCCT<br>ATTGGG/<br>TTAATAT            |
| 75.              | K0124E12-3 | Wbscr5         | Williams-<br>Beuren<br>syndrome<br>chromosome<br>region 5<br>homolog<br>(human) | K0124E12      | Mm.23955      | Chromosome 5                             | AAATGTC/<br>GGTTCCC<br>TCAGAGC<br>TGTGTAC<br>TGTAGC                                  |
| 76.              | K0649H05-3 | F730038I15Rik  | RIKEN cDNA<br>F730038I15<br>gene                                                | K0649H05      | Mm.268680     | No Chromosome location<br>info available | CCAGAG/<br>TCCATCA/<br>TTGCCCC<br>ACCTCGC<br>TAGAAG/<br>AGGAAA                       |
| 77.              | K0154C05-3 | D230024O04     | hypothetical<br>protein<br>D230024O04                                           | K0154C05      | Mm.90241      | Chromosome 6                             | GACAGG/<br>ATGTTTA<br>ACATAC/<br>TGATGA/<br>AATATC/<br>TGAAGA/<br>ACTCTC/<br>CCTGTTC |
| 78.              | C0182E05-3 | Hmox1          | heme<br>oxygenase<br>(decycling) 1                                              | C0182E05      | Mm.230635     | Chromosome 8                             | CAGTGG/<br>GGAATT<br>CATGTA/<br>AATAC                                                |
| 79.              | L0823E04-3 |                | transcribed<br>sequence with<br>weak similarity                                 | L0823E04      | Mm.270136     | Chromosome 7                             | GACAGG<br>CCATATC<br>TAAGGA                                                          |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                    | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A] | 60me<br>SEQUENCE                                             |
|------------------|------------|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|----------------|--------------------------|--------------------------------------------------------------|
|                  |            |                | to protein<br>pir:T26134<br>(C.elegans)<br>T26134<br>hypothetical<br>protein<br>W04A4.5 -<br>Caenorhabditis<br>elegans<br>hypothetical<br>protein<br>9830126M18 |               |                |                          | ACCTCAT<br>AAGTCTC<br>AAAGAA                                 |
| 80.              | K0130E05-3 | 9830126M18     |                                                                                                                                                                 | K0130E05      | Mm.266485      | Chromosome 15            | CTCGGAT<br>ATGTTCT<br>TAAGAA<br>CTGTGGA<br>GAACAA<br>AATAA   |
| 81.              | C0908B11-3 | P2ry6          | pyrimidinergic<br>receptor P2Y,<br>G-protein<br>coupled, 6                                                                                                      | C0908B11      | Mm.32929       | Chromosome 7             | CTAAGAC<br>GTGATTT<br>ACTGGTC<br>CATGCTT<br>CATTCAG<br>CAGGA |
| 82.              | K0438A08-3 | Ccl2           | chemokine (C-<br>C motif) ligand<br>2                                                                                                                           | K0438A08      | Mm.145         | Chromosome 11            | TCCCTCT<br>GAATCCA<br>CAACACT<br>ATGTATG<br>ATGAATT<br>AAAGA |
| 83.              | H3082C12-3 | Spp1           | secreted<br>phosphoprotein<br>1                                                                                                                                 | H3082C12      | Mm.288474      | Chromosome 5             | TTCTCAG<br>GTGGATA<br>TATGTAC<br>AGAGAG<br>TATTTTG<br>CTTAGC |
| 84.              | H3014A12-3 | Capg           | capping protein<br>(actin filament),<br>gelsolin-like                                                                                                           | H3014A12      | Mm.18626       | Chromosome 6             | CTGACCA<br>GGCTGAC<br>GCCCTTT<br>TCTGAAC<br>AATTCCA<br>ACTGC |
| 85.              | H3089C11-3 | BG070621       | ESTs<br>BG070621                                                                                                                                                | H3089C11      | Mm.173282      | Chromosome 4             | GATACCT<br>TATCTTT<br>AACAGC<br>ATGCAG<br>GAAATG<br>ACAGA    |
| 86.              | X67783.1   | Vcam1          | vascular cell<br>adhesion<br>molecule 1                                                                                                                         | X67783        | Mm.76649       | Chromosome 3             | GTTTGAC<br>GACATT<br>TAAAAC<br>ATCCTTA<br>TGTTTAT<br>GCCCCG  |
| 87.              | J0509D03-3 |                | AU018874<br>Mouse eight-<br>cell stage<br>embryo cDNA<br>Mus musculus<br>cDNA clone<br>J0509D03 3',<br>MRNA<br>sequence                                         | J0509D03      |                | Chromosome 13            | CTCTGA<br>AATAAA<br>TGTGAT<br>TATAGT<br>AGTCTT<br>TTAGA      |
| 88.              | H3055A11-5 |                | UNKNOWN:<br>Similar to<br>Homo sapiens<br>KIAA1363<br>protein<br>(KIAA1363),                                                                                    | H3055A11      | Data not found | Chromosome 3             | GGCAAC<br>ACTTTG<br>GCCATG<br>GAACAA<br>ACTTCA<br>TGTAAG     |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                        | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------|-------------|----------------|-----------------------------------------------------|---------------|---------------|------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 89.              | C0455A05-3  | AW413625       | mRNA<br>expressed<br>sequence<br>AW413625           | C0455A05      | Mm.1643       | Chromosome 19                            | ACTTCAT.<br>TTCACAA<br>GAGGGCT<br>AAGATAC<br>ACAATTT<br>CAGTGTG<br>CACCTCT<br>TCCAGCC<br>CCAGGA1<br>TCTAGAA<br>AGGCTAC<br>GCCTG<br>CGTCAGT<br>CACTCAA<br>GTGGTGC<br>GTAAAG1<br>CCAAATC<br>ACCTGT<br>CGAATGA/<br>TGCATCT<br>AGACCA1<br>GAGTTCC<br>GTTTGCT<br>GGAAAG1<br>CCGGCGC<br>TAGTTTC<br>TATTTAG<br>AACTCG1<br>ATATGTA<br>TCTTT<br>CAAGCTC<br>GAGCCTC<br>CTTCAA1<br>TGAATC1<br>AACATT1<br>ACACT<br>CAATCC1<br>CAACTAC<br>GTGTTG1<br>GTTTCA1<br>CATTAAT<br>ATGGG<br>TCCCACC<br>TGATGA1<br>AGCCAA1<br>CCTTAGC<br>TCCATA1<br>TATTCA<br>AAGAAA<br>CCACTTC<br>TGTGTA1<br>TATTTA1<br>AGATAA<br>GCATGC<br>TTTGGG1<br>GCTTCA1<br>GCGCTC1<br>AAAGGA<br>TGTTTCC<br>TATCAA<br>TTTCAT1<br>TAATAT1<br>GGGAGA<br>TAAGTG<br>CACTGT1 |
| 90.              | NM_019732.1 | Runx3          | runt related<br>transcription<br>factor 3           | NM_019732     | Mm.247493     | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 91.              | L0008A03-3  | AW546412       | ESTs<br>AW546412                                    | L0008A03      | Mm.182599     | Chromosome 16                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 92.              | K0329C10-3  | Thbs1          | thrombospondi<br>n 1                                | K0329C10      | Mm.4159       | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 93.              | H3115H03-3  | BC019206       | cDNA sequence<br>BC019206                           | H3115H03      | Mm.259061     | Chromosome 10                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 94.              | C0643F09-3  | Usp18          | ubiquitin<br>specific<br>protease 18                | C0643F09      | Mm.27498      | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 95.              | X84046.1    | Hgf            | hepatocyte<br>growth factor                         | X84046        | Mm.267078     | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 96.              | L0236C05-3  | Aldh1b1        | aldehyde<br>dehydrogenase<br>1 family,<br>member B1 | L0236C05      | Mm.24457      | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 97.              | H3055E08-3  | Mcoln2         | mucolipin 2                                         | H3055E08      | Mm.116862     | Chromosome 3                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 98.              | H3009F12-3  | BG063639       | ESTs<br>BG063639                                    | H3009F12      | Mm.196869     | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 99.              | J0208G12-3  | Cxcl1          | chemokine (C-<br>X-C motif)<br>ligand 1             | J0208G12      | Mm.21013      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60me<br>SEQUENCE                                                                                                                                                                                |
|------------------|------------|----------------|------------------------------------------------------|---------------|---------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 100.             | K0300C11-3 | 9130025P16Rik  | RIKEN cDNA<br>9130025P16<br>gene                     | K0300C11      | Mm.153315     | Chromosome 1             | TAGAAG<br>AAGTGAC<br>TTTTCAT/<br>ACTTAAA<br>GAGTTCC<br>GCCTCTG<br>CTCAG                                                                                                                         |
| 101.             | H3104F03-5 | Krt1-18        | keratin complex<br>1, acidic, gene<br>18             | H3104F03      | Mm.22479      | Chromosome 15            | CAAGGTC<br>AGCCTGC<br>CTGAGAA/<br>GAGACTC<br>AGCAAAA/<br>GGGAAC/<br>GCATGTC<br>ATTCATG<br>CCCCTTA<br>GCAAAGTC<br>CAAAGTT<br>TGAGC                                                               |
| 102.             | L0858D08-3 | Trim2          | tripartite motif<br>protein 2                        | L0858D08      | Mm.44876      | Chromosome 3             | TGCTCCA<br>TGAAACI<br>GACGTAC<br>CCCTGAA/<br>ATTCTA<br>GGAAG<br>TGTACAA/<br>ACTCACC<br>GTGAAG/<br>TGATTGI<br>CTTGTA<br>AGCAC                                                                    |
| 103.             | L0508H09-3 | BY564994       | EST BY564994                                         | L0508H09      | Mm.290934     | Chromosome 12            | TTTTGCA<br>TCGAGTC<br>GCATTG/<br>TAAAACT<br>ATTTGAA/<br>TTCCAT<br>CAAACA/<br>ACAGGG.<br>GTAAAA/<br>TCAACTC<br>AGTTATC<br>CATAGC/<br>TCAGCA/<br>GCGATT/<br>ATCCTAJ<br>CCTACA/<br>AGGAGT<br>GGTGA |
| 104.             | L0701G07-3 | BM194833       | ESTs<br>BM194833                                     | L0701G07      | Mm.221788     | Chromosome 2             | CATGTG/<br>TCATGG/<br>AATAGT/<br>GAATCT/<br>GGTTAG.<br>AAAGAC<br>GTCTCA/<br>CTGGGA/<br>AACTGG<br>GAAAAG<br>GACCAA<br>AGATCA/<br>TGAATC.<br>AAAAGA<br>TGGTGT<br>GAATAT<br>AAGAGT                 |
| 105.             | K0102A10-3 | E430025L02Rik  | RIKEN cDNA<br>E430025L02<br>gene                     | K0102A10      | Mm.33498      | Chromosome 16            |                                                                                                                                                                                                 |
| 106.             | C0190H11-3 | Spn            | sialophorin                                          | C0190H11      | Mm.87180      | Chromosome 7             |                                                                                                                                                                                                 |
| 107.             | L0514A11-3 | 2810457I06Rik  | RIKEN cDNA<br>2810457I06<br>gene                     | L0514A11      | Mm.133615     | Chromosome 9             |                                                                                                                                                                                                 |
| 108.             | J0911E11-3 | Nefl           | neurofilament,<br>light<br>polypeptide               | J0911E11      | Mm.1956       | Chromosome 14            |                                                                                                                                                                                                 |
| 109.             | K0647E02-3 | Def6           | differentially<br>expressed in<br>FDCP 6             | K0647E02      | Mm.60230      | Chromosome 17            |                                                                                                                                                                                                 |
| 110.             | H3091E09-3 | Eif1a          | eukaryotic<br>translation<br>initiation factor<br>1A | H3091E09      | Mm.143141     | Chromosome Un            |                                                                                                                                                                                                 |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                           |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 111.             | AF286725.1 | Pdgfr          | platelet-derived<br>growth factor,<br>C polypeptide                                                                                                  | AF286725      | Mm.40268      | Chromosome 3                             | CAGGTGA<br>AAAGGA/<br>ATATCAG<br>AGATTTG<br>TGATGAG<br>TTCCATC<br>CCCCGA<br>CAGTCCT<br>AAAGGTC<br>AAGCTGC<br>GCAATTA<br>GAGGGA/<br>ACTAATT<br>AGAGGA/<br>CTTATAT<br>GGCAGGC<br>TAGTAAA<br>TCATTTC<br>GAGGA<br>GCATGAC<br>TAGGTGA/<br>TTCACCT<br>ATGCTGT<br>AGTTCTC<br>CTATG |
| 112.             | D31942.1   | Osm            | oncostatin M                                                                                                                                         | D31942        | 18413         | Chromosome 11                            | ATCGTCT<br>TTATGAC<br>GCTATGT<br>TGGCAGC<br>TATTTGA<br>AAAGTG<br>CTGTTCC<br>GGGTTTI<br>ATGTCAC<br>GTGGTTC<br>TCAGGA/<br>GGGAAA<br>GTGCAA/<br>AATATA/<br>TCAAAC/<br>TCTGAAC<br>AGGGCA<br>AGTATA/<br>CTTGTC<br>TGGGGG/<br>ATATCT/<br>TGAAAA<br>ATTTCCT/<br>CAAGA              |
| 113.             | L0046B04-3 | Alcam          | activated<br>leukocyte cell<br>adhesion<br>molecule                                                                                                  | L0046B04      | Mm.2877       | Chromosome 16                            |                                                                                                                                                                                                                                                                             |
| 114.             | K0131D09-3 | LOC217304      | similar to<br>triggering<br>receptor<br>expressed on<br>myeloid cells 5<br>(LOC217304),<br>mRNA                                                      | K0131D09      | Mm.297591     | Chromosome 11                            |                                                                                                                                                                                                                                                                             |
| 115.             | H3024C07-3 | Hexa           | hexosaminidase<br>A                                                                                                                                  | H3024C07      | Mm.2284       | Chromosome 9                             |                                                                                                                                                                                                                                                                             |
| 116.             | L0251A07-3 | B4galt1        | UDP-<br>Gal:betaGlcNA<br>c beta 1,4-<br>galactosyltransf<br>erose,<br>polypeptide 1                                                                  | L0251A07      | Mm.15622      | Chromosome 4                             |                                                                                                                                                                                                                                                                             |
| 117.             | C0612G04-3 | Grip1          | glutamate<br>receptor<br>interacting<br>protein 1                                                                                                    | C0612G04      | Mm.196692     | Chromosome 10                            |                                                                                                                                                                                                                                                                             |
| 118.             | C0357B04-3 |                | C0357B04-3<br>NIA Mouse<br>Undifferentiate<br>d ES Cell<br>cDNA Library<br>(Short) Mus<br>musculus<br>cDNA clone<br>C0357B04 3',<br>MRNA<br>sequence | C0357B04      |               | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                             |
| 119.             | L0529E02-3 | Egfl3          | EGF-like-<br>domain,<br>multiple 3                                                                                                                   | L0529E02      | Mm.29268      | Chromosome 4                             | CAACTG/<br>CTGGAA<br>GTCCAG<br>ATTTAT/<br>GTATTT/<br>CATCT<br>CCTTCC/<br>TTTGCC/<br>TGGAAA<br>GAGATG<br>TACTCC                                                                                                                                                              |
| 120.             | L0218E05-3 | Dnase2a        | deoxyribonucle<br>ase II alpha                                                                                                                       | L0218E05      | Mm.220988     | Chromosome 8                             |                                                                                                                                                                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                                                        | CLONE<br>NAME | UG<br>CLUSTER | CHR LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------|------------|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 121.             | H3074C12-3 | Dutp           | deoxyuridine<br>triphosphatase                                                                                                                                                                                                      | H3074C12      | Mm.173383     | Chromosome 2                             | GTTGG<br>TAGGTGA<br>GGAATCT<br>TAAGGTC<br>ATAGGAT<br>TTATATG<br>AATGG<br>ATGACTT<br>TGCTTGG<br>AGAAGAA<br>TCTTTAC<br>CAGCTTC<br>CTTTTT<br>CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG |
| 122.             | H3072F09-3 | Icsbp1         | interferon<br>consensus<br>sequence<br>binding protein<br>1                                                                                                                                                                         | H3072F09      | Mm.249937     | Chromosome 8                             | ATGACTT<br>TGCTTGG<br>AGAAGAA<br>TCTTTAC<br>CAGCTTC<br>CTTTTT<br>CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG                                                                          |
| 123.             | C0829F05-3 | 4632404H22Rik  | RIKEN cDNA<br>4632404H22<br>gene                                                                                                                                                                                                    | C0829F05      | Mm.28559      | Chromosome X                             | CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG                                                                                                                                           |
| 124.             | L0063A12-3 |                | similar to<br>ubiquitin-<br>conjugating<br>enzyme UBCi<br>(LOC245350),<br>mRNA                                                                                                                                                      | L0063A12      | Mm.38094      | Chromosome X                             | CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG                                                                                                                                           |
| 125.             | C0143E09-3 | 6330548O06Rik  | RIKEN cDNA<br>6330548O06<br>gene                                                                                                                                                                                                    | C0143E09      | Mm.41694      | Chromosome 5                             | CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG                                                                                                                                           |
| 126.             | K0127G03-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_000072.<br>1 (H.sapiens)<br>beige protein<br>homolog;<br>Lysosomal<br>trafficking<br>regulator<br>[Homo sapiens]<br>lysosomal<br>membrane<br>glycoprotein 2 | K0127G03      | Mm.32947      | Chromosome 14                            | CCGGGGT<br>AAGTTGT<br>TCCTGGG<br>TTTCCCA<br>TTTGTTT<br>GCCCT<br>GGAAGA<br>TAAATAG<br>CTGTGGT<br>TTGGAAC<br>GTAGCTT<br>GACACA<br>CCAGGTT<br>GCGGACT<br>ATAATAA<br>TGTATTG<br>AGGAAA<br>GCGGAG<br>TGCATGC<br>ATTTCTA<br>GCTCACT<br>CAAGGCT<br>GCACTGC<br>AAGAAG                                                                                                                                           |
| 127.             | H3109D03-3 | Lamp2          | lysosomal<br>membrane<br>glycoprotein 2                                                                                                                                                                                             | H3109D03      | Mm.486        | Chromosome X                             | TTAACCT<br>GTGCAAC<br>TAATGTC<br>AAGGAC<br>TTCTACA<br>AAGACT<br>TCCCCAC<br>ATAAGG<br>GGAGCT<br>GATCCCC<br>TAAGAA<br>CCCAAA<br>ATAGGT<br>CCGATT<br>GGAGCA<br>GTGGAA<br>GAGTTT<br>AGTAGA<br>AGTAGT<br>CAGTAT<br>ATAAAT<br>TTGACA<br>ATCTTG                                                                                                                                                                |
| 128.             | J0034B02-3 | Dhx16          | DEAH (Asp-<br>Glu-Ala-His)<br>box polypeptide<br>16                                                                                                                                                                                 | J0034B02      | Mm.5624       | Chromosome 17                            | TTAACCT<br>GTGCAAC<br>TAATGTC<br>AAGGAC<br>TTCTACA<br>AAGACT<br>TCCCCAC<br>ATAAGG<br>GGAGCT<br>GATCCCC<br>TAAGAA<br>CCCAAA<br>ATAGGT<br>CCGATT<br>GGAGCA<br>GTGGAA<br>GAGTTT<br>AGTAGA<br>AGTAGT<br>CAGTAT<br>ATAAAT<br>TTGACA<br>ATCTTG                                                                                                                                                                |
| 129.             | K0428C07-3 | Plcb3          | phospholipase<br>C, beta 3                                                                                                                                                                                                          | K0428C07      | Mm.6888       | Chromosome 19                            | TTAACCT<br>GTGCAAC<br>TAATGTC<br>AAGGAC<br>TTCTACA<br>AAGACT<br>TCCCCAC<br>ATAAGG<br>GGAGCT<br>GATCCCC<br>TAAGAA<br>CCCAAA<br>ATAGGT<br>CCGATT<br>GGAGCA<br>GTGGAA<br>GAGTTT<br>AGTAGA<br>AGTAGT<br>CAGTAT<br>ATAAAT<br>TTGACA<br>ATCTTG                                                                                                                                                                |
| 130.             | K0119F10-3 | Ccl9           | chemokine (C-<br>C motif) ligand<br>9                                                                                                                                                                                               | K0119F10      | Mm.2271       | No Chromosome location<br>info available | TTAACCT<br>GTGCAAC<br>TAATGTC<br>AAGGAC<br>TTCTACA<br>AAGACT<br>TCCCCAC<br>ATAAGG<br>GGAGCT<br>GATCCCC<br>TAAGAA<br>CCCAAA<br>ATAGGT<br>CCGATT<br>GGAGCA<br>GTGGAA<br>GAGTTT<br>AGTAGA<br>AGTAGT<br>CAGTAT<br>ATAAAT<br>TTGACA<br>ATCTTG                                                                                                                                                                |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                      |
|------------------|------------|----------------|----------------------------------------------------------------------|---------------|---------------|------------------------------------------|-----------------------------------------------------------------------|
| 131.             | J0046B07-3 | Tuba4          | tubulin, alpha 4                                                     | J0046B07      | Mm.1155       | Chromosome 1                             | CAGCC<br>ACCGCTA<br>GAGCCTG<br>CTGTGTT<br>GCAAAAT<br>TCGAAAT<br>AGTCT |
| 132.             | C0117E11-3 | Neu1           | neuraminidase<br>1                                                   | C0117E11      | Mm.8856       | Chromosome 17                            | TGAACTC<br>CTTTTGC<br>TCTCATC<br>GGGAAGT<br>TCGTTAT<br>TAACA          |
| 133.             | C0101C01-3 | Sdc1           | syndecan 1                                                           | C0101C01      | Mm.2580       | No Chromosome location<br>info available | GTCTGTT<br>GGAATGC<br>AGTAATT<br>CTCTAGC<br>CTTGACC<br>GTCAC          |
| 134.             | K0245A03-3 | 9130012B15Rik  | RIKEN cDNA<br>9130012B15<br>gene                                     | K0245A03      | Mm.35104      | No Chromosome location<br>info available | CCAGCCT<br>AGATTTT<br>ACCTTTT<br>AAGAGAC<br>ATTCTAA<br>ATAAA          |
| 135.             | H3109A02-3 | Fcrlg          | Fc receptor,<br>IgE, high<br>affinity I,<br>gamma<br>polypeptide     | H3109A02      | Mm.22673      | Chromosome 1                             | CACCTCT<br>TTTGAAG<br>GCTGACC<br>TCCCATA<br>TGCTAGC<br>CTTTA          |
| 136.             | L0819C05-3 | Mapk8ip        | mitogen<br>activated<br>protein kinase 8<br>interacting<br>protein   | L0819C05      | Mm.2720       | Chromosome 2                             | CTGAGCT<br>CTGAGCC<br>CACCTCC<br>GACTTTC<br>AAGGAA<br>CAACGT          |
| 137.             | U77083.1   | Anpep          | alanyl<br>(membrane)<br>aminopeptidase                               | U77083        | Mm.4487       | Chromosome 7                             | AGAACA<br>TTAGTTC<br>TTCTGAC<br>ACTTGTC<br>TATGACA<br>TTACTA          |
| 138.             | C0164B01-3 | Tnfaip2        | tumor necrosis<br>factor, alpha-<br>induced protein<br>2             | C0164B01      | Mm.4348       | Chromosome 12                            | ATGTGTC<br>CAGGAC<br>TCCAGAC<br>CTTTTTC<br>AGCTTG<br>AAACAG           |
| 139.             | H3085G03-3 | Cyba           | cytochrome b-<br>245, alpha<br>polypeptide                           | H3085G03      | Mm.448        | Chromosome 8                             | ACGTTTC<br>AGTGGT<br>GGCGCC<br>TATCGC<br>GTGTGC<br>TGTCT              |
| 140.             | H3074F04-3 | Abcc3          | ATP-binding<br>cassette, sub-<br>family C<br>(CFTR/MRP),<br>member 3 | H3074F04      | Mm.23942      | Chromosome 11                            | TTTTTTA<br>GCAAAT<br>CACAGT<br>GAGGAA<br>GTTAGA<br>CAGCC              |
| 141.             | H3145E02-3 | Wbp1           | WW domain<br>binding protein<br>1                                    | H3145E02      | Mm.1109       | Chromosome 6                             | GTGCTA<br>ACTCAC<br>AGACAT<br>AGGAGC<br>ATCTCA                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                     |
|------------------|------------|----------------|----------------------------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------|
| 142.             | K0609F07-3 | Cd53           | CD53 antigen                                                         | K0609F07      | Mm.2692       | Chromosome 3             | GAGACA<br>GAGGTCC<br>TTAAATG<br>TCTCCTA<br>CTGTCAA<br>ATTCTA<br>CTAAA |
| 143.             | K0205H04-3 | 9830148O20Rik  | RIKEN cDNA<br>9830148O20<br>gene                                     | K0205H04      | Mm.21630      | Chromosome 9             | CTTCTAG<br>TTCTGCA<br>TCATCGT<br>AAGGAG<br>AACTATT<br>CGAAT           |
| 144.             | H3095H04-3 | 2410002I16Rik  | RIKEN cDNA<br>2410002I16<br>gene                                     | H3095H04      | Mm.17537      | Chromosome 18            | ACTTATT<br>CTTGCCT<br>CCACCCC<br>AAACAG<br>ATTAATA<br>ATGTG           |
| 145.             | C0623H08-3 | Tm7sf1         | transmembrane<br>7 superfamily<br>member 1                           | C0623H08      | Mm.1585       | Chromosome 13            | TACAGTA<br>GCAAGC<br>TCCATTT<br>AATAAAA<br>CAGCAT<br>TCAGC            |
| 146.             | L0242F05-3 | 2700088M22Rik  | RIKEN cDNA<br>2700088M22<br>gene                                     | L0242F05      | Mm.103104     | Chromosome 15            | TTATTTA<br>ATCTTAC<br>TAACCT<br>GACCTG<br>CACTGG<br>TAGAC             |
| 147.             | C0177F02-3 | Sdc3           | syndecan 3                                                           | C0177F02      | Mm.206536     | Chromosome 4             | CCTGTCC<br>TTCATGC<br>AACTTA<br>GAGAAG<br>AGAGGG<br>ATGGAT            |
| 148.             | L0803B02-3 | Ppp1r9a        | protein<br>phosphatase 1,<br>regulatory<br>(inhibitor)<br>subunit 9A | L0803B02      | Mm.156600     | Chromosome 6             | AAAGGG<br>GAGTAT<br>GTTGCA<br>TATACT<br>TCCTCC<br>GTTTAT              |
| 149.             | H3061D01-3 | BB172728       | ESTs<br>BB172728                                                     | H3061D01      | Mm.254385     | Chromosome 3             | TATCCG<br>TCTATG<br>TAGGAC<br>GTCGAA<br>GGAAAG<br>CAACAG              |
| 150.             | L0259D11-3 | C1qb           | complement<br>component 1, q<br>subcomponent,<br>beta<br>polypeptide | L0259D11      | Mm.2570       | Chromosome 4             | CTGCTT<br>TGACAT<br>CGTAAT<br>GGTCAA<br>ACCTAT<br>ACCAT               |
| 151.             | H3011D10-3 | Lcp1           | lymphocyte<br>cytosolic<br>protein 1                                 | H3011D10      | Mm.153911     | Chromosome 14            | AACAAA<br>ACAGTA<br>TTGAAT<br>CCACTA<br>GCAATT<br>GAGAAC              |
| 152.             | H3052B11-3 | Pctk3          | PCTAIRE-<br>motif protein<br>kinase 3                                | H3052B11      | Mm.28130      | Chromosome 1             | CTGACT<br>TGTCGT<br>AGAGCA<br>CAGAGA<br>ATTTAA                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------|------------|----------------|-------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 153.             | K0413H04-3 | Anxa8          | annexin A8                                      | K0413H04      | Mm.3267       | Chromosome 14            | GTTGTAC<br>GCCTGAA<br>CATGACA<br>CTCTTCT<br>ATTTCGTT<br>TTTCAGA<br>TAAACAT<br>CCTGTGT<br>AAAAAT/<br>GAACTGC<br>AGGAGAC<br>TTGATCT<br>AAACAGC<br>GTAAAGA/<br>TAGACTC<br>GAGTTAA/<br>GCACTCT<br>TTACCAT<br>TTTGG                                                                                                                                                                                                                                                                                      |
| 154.             | H3054F05-3 | Lyzs           | lysozyme                                        | H3054F05      | Mm.45436      | Chromosome 10            | TGTGAA/<br>GTGCATC<br>TTCAACT<br>TGAACCC<br>GGAAGA.<br>GATTCC<br>AGCTGCC<br>AGCAGT.<br>AAGGAG<br>CTGTCTC<br>AGGTGA.<br>AAATGT<br>CCATGTT<br>AGTATG.<br>AAGAGG<br>TATTAAC<br>TGAAAG.<br>GAATAC<br>GTGAAT<br>GCATAG<br>TTTGTAT<br>ATGTTCC<br>AAGTGT<br>TGAAC<br>ACCCAC<br>AGGATA<br>GAAAGG<br>TGACCT<br>ACGCAT<br>TCCTGC/<br>CCCGCT<br>GAGAAC<br>GGAGAG<br>GTGTGT.<br>GAAGCA<br>AATAAC<br>ATATTA<br>ATAAAA<br>GCTGTC<br>AATGGA<br>CTTTCT/<br>TCCCAC<br>TGTGGG<br>TGAAGA<br>TGAGCA<br>ATAGAA<br>GACTGC |
| 155.             | H3060F11-3 | Cybb           | cytochrome b-<br>245, beta<br>polypeptide       | H3060F11      | Mm.200362     | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 156.             | H3012F08-3 | 9430068N19Rik  | RIKEN cDNA<br>9430068N19<br>gene                | H3012F08      | Mm.143819     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 157.             | G0106B08-3 | Abr            | active BCR-<br>related gene                     | G0106B08      | Mm.27923      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 158.             | L0287A12-3 | Tdrkh          | tudor and KH<br>domain<br>containing<br>protein | L0287A12      | Mm.40894      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 159.             | H3083D01-3 | AY007814       | hypothetical<br>protein,<br>12H19.01.T7         | H3083D01      | Mm.160389     | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 160.             | H3131F02-3 | BG074151       | ESTs<br>BG074151                                | H3131F02      | Mm.142524     | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 161.             | C0172H02-3 | Lgals3         | lectin, galactose<br>binding; soluble<br>3      | C0172H02      | Mm.2970       | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 162.             | K0542E07-3 | Cd44           | CD44 antigen                                    | K0542E07      | Mm.24138      | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 163.             | C0450H11-3 | E430019N21Rik  | RIKEN cDNA<br>E430019N21<br>gene                | C0450H11      | Mm.275894     | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                             | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------|------------|----------------|--------------------------------------------------------------------------|---------------|---------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 164.             | K0216A08-3 | Orc5l          | origin<br>recognition<br>complex,<br>subunit 5-like<br>(S. cerevisiae)   | K0216A08      | Mm.566        | Chromosome 5             | TCCTG<br>CTACTCT<br>GATGTTA<br>AACACTG<br>TGCCTGA<br>CATTTAC<br>GACTG<br>TCGACCA<br>TAGGCAC<br>TTCTGGG<br>GGCGCTC<br>GACATAI<br>TTTAT<br>TCTGAAT<br>GCACTGA<br>GATGCAI<br>ATAATGT<br>GTTTTC<br>TGTCTTC<br>GATCCTT<br>CTCCATA<br>GATTTTT<br>TAGTTAA<br>TGTAAC<br>ACACA<br>CTCAGCA<br>CAGAGA<br>ATGAATC<br>CCACTGA<br>TCGTGA<br>TGAATCT<br>CTTTGTT<br>CCCAGC<br>AAGCCA<br>ATAACA<br>CTCATGT<br>GCAAA<br>AAATTG<br>GCATCC<br>GGGGAG<br>TAACCA<br>ATCACC<br>GAATT<br>ACATGA<br>AAGAAT<br>AAGATC<br>TGTCTAC<br>TTCAGA<br>TTACA<br>CACCCC<br>AAATGA<br>ATTGAA<br>TCCTTTC<br>AAGATC<br>ACAGGA<br>AGTGAT<br>ACCATG<br>GCTGTA<br>AACCTC<br>CTGCAA<br>CTACTG<br>AAAGGT<br>GGTTTC<br>GTTTGG<br>GGAGTC<br>GTTGCA |
| 165.             | H3122D03-3 | Pdgfr          | platelet-derived<br>growth factor,<br>C polypeptide                      | H3122D03      | Mm.40268      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 166.             | C0037H07-3 | Il13ra1        | interleukin 13<br>receptor, alpha<br>1                                   | C0037H07      | Mm.24208      | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 167.             | H3054F04-3 | 2610318I15Rik  | RIKEN cDNA<br>2610318I15<br>gene                                         | H3054F04      | Mm.34490      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 168.             | L0908A12-3 | Blnk           | B-cell linker                                                            | L0908A12      | Mm.9749       | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 169.             | G0111E06-3 | Car7           | carbonic<br>anhydrase 7                                                  | G0111E06      | Mm.154804     | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 170.             | L0284B06-3 | Ngfrap1        | nerve growth<br>factor receptor<br>(TNFRSF16)<br>associated<br>protein 1 | L0284B06      | Mm.90787      | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 171.             | K0145G06-3 | Tcfec          | transcription<br>factor EC                                               | K0145G06      | Mm.36217      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 172.             | H3001B08-3 | Lyn            | Yamaguchi<br>sarcoma viral<br>(v-yes-1)<br>oncogene<br>homolog           | H3001B08      | Mm.1834       | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 173.             | G0117F12-3 | Prkcsh         | protein kinase<br>C substrate<br>80K-H                                   | G0117F12      | Mm.214593     | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 174.             | C0903A11-3 | 2510004L01Rik  | RIKEN cDNA<br>2510004L01<br>gene                                         | C0903A11      | Mm.24045      | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                      | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------|------------|----------------|-------------------------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 175.             | L0062C10-3 | Rasa3          | RAS p21<br>protein<br>activator 3                                 | L0062C10      | Mm.18517      | Chromosome 8             | AAACAG<br>TCTATGT<br>TAGGGGC<br>CCCAGGC<br>TCCAAAG<br>ACAGTAT<br>TTTCTCA<br>CTACACA<br>CTTTAGG<br>TAGGTTT<br>CTGAGCC<br>TTTCGAT<br>CACTG<br>AAGTCTA<br>GGAATGC<br>CTCAATG<br>TTGTTCT<br>AATGATA<br>AAATAA<br>GGAAGA<br>GACCTCA<br>AAAAAT<br>TACGACC<br>AATTGCA<br>TATATTC<br>GGATGA<br>ACTGAG<br>CCCTTCT<br>TCTTCAT<br>CAAGCA<br>CACCAT<br>CTGTTC<br>CAAACA<br>GTTCTC<br>GGGACA<br>CATCAT<br>GGAAAA<br>AGGAGT<br>GTTTTGA<br>TGTATT<br>TTGGAA<br>ACCAAC<br>GCTCAG<br>CTCAAT<br>CTCTAA<br>CATCAC<br>AGTCTT<br>TTCATG<br>TTAAT<br>ACTTAA<br>AGACTG<br>ACAGTG<br>CAGTAT<br>GAATGT<br>TTACT<br>TTTCAT<br>AACCGT<br>AGTGAC<br>GATTAT<br>GATTTG<br>AAAAC<br>ATCCAT<br>ATCAAT<br>TATGTA<br>ATGACT<br>AGGGCC |
| 176.             | L0939G09-3 | Cd38           | CD38 antigen                                                      | L0939G09      | Mm.249873     | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 177.             | H3115B07-3 | S100a9         | S100 calcium<br>binding protein<br>A9 (calgranulin<br>B)          | H3115B07      | Mm.2128       | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 178.             | K0608H07-3 | Fyb            | FYN binding<br>protein                                            | K0608H07      | Mm.254240     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 179.             | C0104E07-3 | Tcird1         | T-cell, immune<br>regulator 1                                     | C0104E07      | Mm.19185      | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 180.             | K0431D02-3 | Wisp1          | WNT1<br>inducible<br>signaling<br>pathway protein<br>1            | K0431D02      | Mm.10222      | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 181.             | L0837H10-3 | Igfbp2         | insulin-like<br>growth factor<br>binding protein<br>2             | L0837H10      | Mm.141936     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 182.             | C0159A08-3 | Mta3           | metastasis<br>associated 3                                        | C0159A08      | Mm.18821      | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 183.             | K0649D06-3 | Ms4a6b         | membrane-<br>spanning 4-<br>domains,<br>subfamily A,<br>member 6B | K0649D06      | Mm.29487      | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 184.             | K0609D11-3 | Man1a          | mannosidase 1,<br>alpha                                           | K0609D11      | Mm.117294     | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 185.             | C0907B04-3 | Mcoln3         | mucopolip 3                                                       | C0907B04      | Mm.114683     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------|------------|----------------|------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 186.             | H3020D08-3 | Edem1          | ER degradation<br>enhancer,<br>mannosidase<br>alpha-like 1 | H3020D08      | Mm.21596      | Chromosome 6                             | AAACC<br>CACAAAA<br>AAATGTG/<br>TCGTACG/<br>ATCACGT/<br>GACAAGT<br>AGAAGA<br>CTATCAG/<br>GTGATAA<br>CGTCATT/<br>GACATTA/<br>GACATGG<br>CGATGA<br>GGAGATC<br>CTCTTGT/<br>AATATAC<br>TCCAAAC<br>TTAGAGC<br>TAGGC<br>ACTATTA<br>CTCAGGA<br>GTAGGAA<br>TTTCCTT/<br>ACAGTTT<br>TCAGTA<br>AAAGAG/<br>TATGTCA<br>GTGATAC<br>GCAACTC<br>GTGGTGA<br>GTTCTTG<br>GAGAGA/<br>TGGGGCC<br>AGAAAA/<br>GGATTTT<br>AAAGCA/<br>CACAAAC<br>GTTGTAC<br>GGAAAG/<br>GCTGGG/<br>CAATATC<br>AGAAAA.<br>AGTTGT<br>AGACCA/<br>CACGGA/<br>TGGATG/<br>ATCTACT<br>CAAGGC/<br>GTCTTCT<br>CAGAGC.<br>GGCTTTT<br>TTATTTT<br>TGGAAA.<br>CAATAA.<br>TTTGTA<br>CTTGGC/<br>TCCTTAC<br>GGGACA<br>CACTGT/<br>TGCCAG/<br>GAATCT<br>ACTTAT/<br>AGGACA<br>GAAGCC<br>AAGAAA<br>AGAAAG |
| 187.             | J0039F05-3 | Gdf3           | growth<br>differentiation<br>factor 3                      | J0039F05      | Mm.4213       | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 188.             | C0906C11-3 | BM218094       | ESTs<br>BM218094                                           | C0906C11      | Mm.212279     | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 189.             | L0266E10-3 | B930060C03     | hypothetical<br>protein<br>B930060C03                      | L0266E10      | Mm.89568      | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 190.             | H3060D11-3 | Mll5           | myeloid/lymph<br>oid or mixed-<br>lineage<br>leukemia 5    | H3060D11      | Mm.10878      | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 191.             | L0062E01-3 | Tnc            | tenascin C                                                 | L0062E01      | Mm.980        | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 192.             | K0132G08-3 | AI662270       | expressed<br>sequence<br>AI662270                          | K0132G08      | Mm.37773      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 193.             | H3114D08-3 | Arpc3          | actin related<br>protein 2/3<br>complex,<br>subunit 3      | H3114D08      | Mm.24498      | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 194.             | C0649E02-3 | Unc93b         | unc-93<br>homolog B (C.<br>elegans)                        | C0649E02      | Mm.28406      | Chromosome 19                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 195.             | L0293H10-3 | 2510048K03Rik  | RIKEN cDNA<br>2510048K03<br>gene                           | L0293H10      | Mm.39856      | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 196.             | H3024C03-3 | 1110008B24Rik  | RIKEN cDNA<br>1110008B24<br>gene                           | H3024C03      | Mm.275813     | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                | CLONE<br>NAME | UG<br>CLUSTER | CHR LOCATION<br>PENG [A] | 60me<br>SEQUEN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------|------------|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 197.             | H3055G02-3 | Ctsc           | cathepsin C                                                                                                                                                 | H3055G02      | Mm.684        | Chromosome 7             | GAGCGCG<br>TAGTTCAC<br>ACAAGTA<br>TCAATGA<br>GCTGTGT<br>ATCAAGT<br>TGTTTC<br>CATGAAT<br>AAACCTA<br>CAAAGCA<br>GTCTCTT<br>GTGAGGT<br>GAACCC<br>CCTGTCT<br>GGAGATT<br>TCATAAG<br>AATCACT<br>GTAACCT<br>GAGGAA<br>AAGTAAA<br>CAAAGGA<br>AAGTTAG<br>AACTCCT<br>TAAGAAA<br>GTCTTCC<br>TCGGAAC<br>CCTTAAG<br>GTGATAT<br>AAGATCC<br>TAAGAAC<br>CAGCAA<br>TTAGTGG<br>AACCTAT<br>TTAACT<br>GTCTTAA<br>CCATAAA<br>GAGAA<br>TTTGTGA<br>CCTGACT<br>CTCCACA<br>TTTCTAT<br>CATGTAA<br>CAATA<br>AGACTTC<br>AAGGCTT<br>ACAATT<br>AAAACCC<br>TCCCACC<br>TCTTGAC<br>TAATAAA<br>ACTGTGC<br>ACTTGG<br>TACTGA<br>AAAGAC<br>GGCTGG |
| 198.             | K0518A04-3 | BM238476       | ESTs<br>BM238476                                                                                                                                            | K0518A04      | Mm.217227     | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 199.             | K0128H01-3 | Parvg          | parvin, gamma                                                                                                                                               | K0128H01      | Mm.202348     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 200.             | K0649F04-3 | Ccr2           | chemokine (C-<br>C) receptor 2                                                                                                                              | K0649F04      | Mm.6272       | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 201.             | K0603E03-3 | Vav1           | vav 1 oncogene                                                                                                                                              | K0603E03      | Mm.254859     | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 202.             | K0649A02-3 | Stat1          | signal<br>transducer and<br>activator of<br>transcription 1                                                                                                 | K0649A02      | Mm.8249       | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 203.             | H3013D11-3 | Mt2            | metallothionein<br>2                                                                                                                                        | H3013D11      | Mm.147226     | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 204.             | H3013B02-3 | Atp6v1b2       | ATPase, H+<br>transporting,<br>V1 subunit B,<br>isoform 2                                                                                                   | H3013B02      | Mm.10727      | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 205.             | L0541H09-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | L0541H09      | Mm.221768     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 206.             | K0516E03-3 |                | Mus musculus<br>12 days embryo<br>embryonic<br>body between<br>diaphragm<br>region and neck                                                                 | K0516E03      | Mm.214742     | Chromosome 10            | AGGTTA<br>ATATTCT<br>AACATG<br>ACAACT<br>AAACCG<br>CCACCA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                                       | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60me SEQUEL                                                                                                                                                                                                                                                                                                                                                                                            |
|------------|------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            |            |               | cDNA, RIKEN full-length enriched library, clone:9430012 B12 product:unknown EST, full insert sequence. urokinase plasminogen activator receptor |            |            |                                       |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 207.       | H3034A10-3 | Plaur         |                                                                                                                                                 | H3034A10   | Mm.1359    | Chromosome 7                          | CCTCGTG<br>CTTCTTTC<br>CTCAGTT<br>CATGAAC<br>AAGAGAA<br>GAACAAC<br>AATAGCA<br>ATCAAAC<br>GATGTGA<br>AGATGCG<br>ATCATCA<br>AATGCC<br>TCTCTGG<br>ATCAGTA<br>CAAAAGC<br>AGAGGGT<br>AAAGCAC<br>AGTAAT<br>TGGAATC<br>AGAATGA<br>CTCGAGC<br>TAGAGGT<br>GTCATCT<br>ATTCAG<br>ATAGAAC<br>GTAGGA/<br>CAGGCA/<br>AAAAATG/<br>CAGTCCA<br>ATCATGC<br>AGATGGC<br>AAGTACT<br>GTTCTTG<br>CTGGATC<br>GCAGAA<br>ACTGTCT |
| 208.       | C0910G05-3 | BM218419      | ESTs<br>BM218419                                                                                                                                | C0910G05   | Mm.217839  | Chromosome 10                         |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 209.       | C0262H12-3 | Msh2          | mutS homolog 2 (E. coli)                                                                                                                        | C0262H12   | Mm.4619    | Chromosome 17                         |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 210.       | H3078C11-3 | BG069620      | ESTs<br>BG069620                                                                                                                                | H3078C11   | Mm.173427  | Chromosome 2                          |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 211.       | L0926H09-3 | 6030440G05Rik | RIKEN cDNA 6030440G05 gene                                                                                                                      | L0926H09   | Mm.27789   | Chromosome 6                          |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 212.       | J0076H03-3 |               | C80125 Mouse 3.5-dpc blastocyst cDNA Mus musculus cDNA clone J0076H03 3', MRNA sequence transcribed                                             | J0076H03   |            | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                        |
| 213.       | L0817B08-3 |               | sequence with strong similarity to protein sp:P00722 (E. coli) BGAL_ECOLI Beta-galactosidase (Lactase)                                          | L0817B08   | Mm.221816  | Chromosome 18: not placed             | AGGAAA<br>CGGTAG<br>ACATCTC<br>CTCAAT<br>GATTGC<br>GTGAAA                                                                                                                                                                                                                                                                                                                                              |
| 214.       | H3065D11-3 | Crnk1l        | Crn, crooked neck-like 1 (Drosophila)                                                                                                           | H3065D11   | Mm.273506  | Chromosome 2                          | GTTTTTC<br>TTGGAC<br>AATTGT<br>ATGGAT<br>TTGCAG<br>GAGAC<br>TGGGAC                                                                                                                                                                                                                                                                                                                                     |
| 215.       | H3157E02-3 | 5630401J11Rik | RIKEN cDNA                                                                                                                                      | H3157E02   | Mm.21104   | Chromosome 17                         |                                                                                                                                                                                                                                                                                                                                                                                                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60me<br>SEQUEN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------|------------|----------------|------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 5630401J11<br>gene                                         |               |               |                                          | AAGCGAC<br>AGAAAAT<br>GAAACAA<br>GATTGCT<br>ACAATTA<br>TCCATTAT<br>ATACAAC<br>AAGAAAA<br>CAGAAAA<br>CCCTTAG<br>ATCAGGG<br>ATTCAAC<br>TTCTAGG<br>TGGCAAG<br>GTAAATT<br>TCCATTTC<br>TCTGTG<br>CCATATG<br>CAGTCGT<br>AACTGCA<br>ACTGAAA<br>ATTATAT<br>GCCAGC<br>GGGCCAT<br>TAAAGAT<br>GAGAGAC<br>CTAGCAT<br>AATTTTC<br>TATTGAG<br>GAAAGGC<br>ATTCAGA<br>GATGGTA<br>TCAGACT<br>AGCACAC<br>ACCCA<br>TAAGGTC<br>CTCCAGT<br>TTCAGTT<br>AATAGTA<br>TTGCCCC<br>GCAAC<br>CCACCA<br>GAAAAA<br>ATGTGTA<br>TAGGTGT<br>CTATGTC<br>ATTGGC<br>GCACAA<br>GAGTCA<br>ATTAAGC<br>ATCATTI<br>CATATA<br>GCAGAG<br>GATTAA<br>ATTAGGC<br>GAAATA<br>GGGCTTC<br>TGTGTAC<br>TAGAGC<br>TGAAGT<br>CTCTAA<br>AATGGG<br>AATATG<br>GTAGGA<br>AGGAAG<br>GTGTAA |
| 216.             | H3007C11-3 | BG063444       | ESTs<br>BG063444                                           | H3007C11      | Mm.182542     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 217.             | K0517E07-3 | C530050H10Rik  | RIKEN cDNA<br>C530050H10<br>gene                           | K0517E07      | Mm.260378     | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 218.             | H3150B11-5 | Ptpn2          | protein tyrosine<br>phosphatase,<br>non-receptor<br>type 2 | H3150B11      | Mm.260433     | Chromosome 18                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 219.             | C0199C01-3 | 9930104E21Rik  | RIKEN cDNA<br>9930104E21<br>gene                           | C0199C01      | Mm.29216      | Chromosome 18                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 220.             | H3063A09-3 | Rassf5         | Ras association<br>(RalGDS/AF-6)<br>domain family<br>5     | H3063A09      | Mm.248291     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 221.             | K0445A07-3 | Hfe            | hemochromatosis                                            | K0445A07      | Mm.2681       | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 222.             | H3123G07-3 | C630007C17Rik  | RIKEN cDNA<br>C630007C17<br>gene                           | H3123G07      | Mm.119383     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 223.             | H3094C03-3 | Baz1a          | bromodomain<br>adjacent to zinc<br>finger domain<br>1A     | H3094C03      | Mm.263733     | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 224.             | L0845H04-3 | BM117070       | ESTs<br>BM117070                                           | L0845H04      | Mm.221860     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 225.             | C0161F01-3 | BC010311       | cDNA sequence<br>BC010311                                  | C0161F01      | Mm.46455      | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 226.             | H3034E07-3 | BG065726       | ESTs                                                       | H3034E07      | Mm.5522       | Chromosome 9                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60me<br>SEQUENCE                                                                                                                                                                    |
|------------------|------------|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | BG065726                                                                                                                                                        |               |               |                                          | AGATGGG<br>GACAATA<br>ATGAAGG<br>GGTAAGA<br>ACCAGAC<br>GGGAAAT<br>CAGCGTT<br>GTTTCCA<br>TGATTTT<br>GAATGAG<br>TATGTG                                                                |
| 227.             | J0419G11-3 | Cldn8          | claudin 8                                                                                                                                                       | J0419G11      | Mm.25836      | Chromosome 16                            | GTAGGAC<br>GAACTGT<br>GGAAGAA<br>GAACATT<br>AATGTGT<br>AATTGAA<br>TCATAGG<br>CATTTAG<br>AGTGTTT<br>GACAATC<br>AAGTTTA<br>CATAGG                                                     |
| 228.             | C0040C08-3 | Cxcr4          | chemokine (C-<br>X-C motif)<br>receptor 4                                                                                                                       | C0040C08      | Mm.1401       | Chromosome 1                             | TTGGAAT<br>GAATGAC<br>GAAATGC<br>AAACTGC<br>CCCGAGT<br>GAATGTC                                                                                                                      |
| 229.             | K0612H02-3 | BM241159       | ESTs<br>BM241159                                                                                                                                                | K0612H02      | Mm.222325     | Chromosome 16                            | CTATCTT<br>TGCTAGA<br>AGAGAA<br>AATGTTA<br>AAAATAC<br>CCTGGCC                                                                                                                       |
| 230.             | J0460B09-3 |                | AU024759<br>Mouse<br>unfertilized egg<br>cDNA Mus<br>musculus<br>cDNA clone<br>J0460B09 3',<br>MRNA<br>sequence                                                 | J0460B09      |               | No Chromosome location<br>info available |                                                                                                                                                                                     |
| 231.             | H3103F07-3 |                | Mus musculus<br>transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus] | H3103F07      | Mm.174026     | Chromosome 10                            |                                                                                                                                                                                     |
| 232.             | H3079H09-3 | BG069769       | ESTs<br>BG069769                                                                                                                                                | H3079H09      | Mm.173446     | Chromosome 9                             | AATCCCT<br>AAAATGC<br>TAGAAA<br>CTGCAT<br>CTCAAAC<br>AGATAC<br>AGACTG<br>AAACCT<br>TACCCA<br>CAGGGG<br>ATAGCA<br>GTCTCA<br>AGAGGA<br>TGTCTG<br>GATATT<br>CTACTTC<br>AAATGA<br>TTGCT |
| 233.             | H3130D06-3 | BG074061       | ESTs<br>BG074061                                                                                                                                                | H3130D06      | Mm.182873     | Chromosome 1                             | ATGGAG<br>TAAACA<br>GACATT<br>AACTAT<br>GTCAGT<br>GTTTCA                                                                                                                            |
| 234.             | H3071D08-3 | Lcp2           | lymphocyte<br>cytosolic<br>protein 2                                                                                                                            | H3071D08      | Mm.1781       | Chromosome 11                            |                                                                                                                                                                                     |
| 235.             | K0218E07-3 |                | Mus musculus<br>10 days neonate<br>olfactory brain<br>cDNA, RIKEN<br>full-length<br>enriched                                                                    | K0218E07      | Mm.216167     | Chromosome 10                            |                                                                                                                                                                                     |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                                                          | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                    | 60me SEQUE                                                                                                                                                                                                                                                         |
|------------|------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 236.       | C0907H07-3 | BM218221      | library,<br>clone:E530016<br>P10<br>product:weakly<br>similar to<br>ONCOGENE<br>TLM [Mus<br>musculus], full<br>insert sequence.<br>ESTs<br>BM218221                | C0907H07   | Mm.221604  | Chromosome 12                            | GAGGCTA<br>AAATAAC<br>AATGCAT<br>GAACTGA<br>GTAATAA<br>GCTCC                                                                                                                                                                                                       |
| 237.       | K0605B09-3 | BM240642      | ESTs<br>BM240642                                                                                                                                                   | K0605B09   | Mm.222320  | Chromosome X                             | AAAGTCGG<br>ATGTCTT<br>TTCTTCTC<br>TAGCTCA<br>AAGATGC<br>CTCAAGT<br>CACTTTT<br>GAAGAA/<br>GTGTGTA<br>TTCCGTG<br>TAGTAAT<br>ATATCT                                                                                                                                  |
| 238.       | C0322F05-3 | Eya3          | eyes absent 3<br>homolog<br>(Drosophila)                                                                                                                           | C0322F05   | Mm.1430    | Chromosome 4                             | TGTAAGA<br>AAGGTA/<br>AAAATAC<br>AATACAC<br>CATATCT<br>ATCGCCC<br>CAGAAA/<br>AGTATGC<br>AAATCAC<br>AGGGAA/<br>AGGGAT/<br>AGCCAA/<br>ACTGAA/<br>GGGAGA/<br>TGTAATT<br>AGGATA/<br>ACTTAGC<br>GACAAC/<br>AAGCTG/<br>TATGGA/<br>CTGTAA/<br>AGAGTG/<br>TTTTGAC<br>GAGTT |
| 239.       | J0004A01-3 | C76123        | ESTs C76123                                                                                                                                                        | J0004A01   | Mm.24905   | Chromosome 15                            | TTTATCA<br>TGGAAA/<br>AGAGAC<br>GGAGAG<br>TGGGTT/<br>ATATGG                                                                                                                                                                                                        |
| 240.       | K0139H06-3 | BM223668      | ESTs<br>BM223668                                                                                                                                                   | K0139H06   | Mm.221718  | Chromosome 3                             | GGAAGT<br>GAACTG<br>AATGTG                                                                                                                                                                                                                                         |
| 241.       | L0941F06-3 | BM120591      | ESTs<br>BM120591                                                                                                                                                   | L0941F06   | Mm.217090  | Chromosome 9                             |                                                                                                                                                                                                                                                                    |
| 242.       | C0300G03-3 | 3021401C12Rik | RIKEN cDNA<br>3021401C12<br>gene                                                                                                                                   | C0300G03   | Mm.102470  | Chromosome 15                            |                                                                                                                                                                                                                                                                    |
| 243.       | C0925E03-3 |               | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | C0925E03   | Mm.217865  | Chromosome 6                             |                                                                                                                                                                                                                                                                    |
| 244.       | H3083B07-5 | BG082983      | ESTs<br>BG082983                                                                                                                                                   | H3083B07   | Mm.203206  | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                    |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                       | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60me SEQUEN                                                                                                                                                                                                                                                                                                          |
|------------|------------|---------------|-----------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 245.       | H3056F01-3 | Gdf9          | growth differentiation factor 9                                                                                 | H3056F01   | Mm.9714    | Chromosome 11                         | GGAAATA/<br>TCAATAAC<br>AAGCCCC<br>AGTGTAG/<br>CAGTGGAG/<br>ATTTGTTA<br>TAAGTCT/<br>TAGAATG/<br>TGTGAA<br>GAAAGTG<br>AATGAAA<br>ATAACAA<br>AAAAAGA<br>TTTCTAG/<br>TTTAGGC/<br>GGTTTTCT/<br>GTTTATC/<br>ATTCTTT/<br>GAAGCAA<br>ATCCATT<br>TGTTGG<br>CTTTTTG/<br>TTATTTT/<br>CAGTTTT/<br>TGTTTAT/<br>CATTTTC/<br>TTACT |
| 246.       | J0259A06-3 | C88243        | EST C88243                                                                                                      | J0259A06   | Mm.249965  | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                      |
| 247.       | C0124B09-3 | BC042513      | cDNA sequence BC042513                                                                                          | C0124B09   | Mm.11186   | Chromosome 11                         |                                                                                                                                                                                                                                                                                                                      |
| 248.       | L0933E02-3 |               | L0933E02-3 NIA Mouse Newborn Kidney cDNA Library (Long) Mus musculus cDNA clone L0933E02 3', MRNA sequence ESTs | L0933E02   |            | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                      |
| 249.       | H3072B12-3 | BG069052      | ESTs BG069052                                                                                                   | H3072B12   | Mm.250102  | Chromosome 9                          | AGTGTTT<br>TTAATTC<br>GGTTGTT<br>TAATATT<br>TATAGTG<br>AATGT<br>TAAAGTA/<br>CTGAAGT/<br>ATGGAA/<br>GCCTTTT<br>TATGGAC<br>TAGCTC<br>GCCTAGT/<br>TCAGCA/<br>TTTGGA/<br>TTAGACC<br>GCATATT<br>CAAGT<br>TCATTTT<br>GTCGTC/<br>GATGTTT<br>TTTCCG<br>GACTTG/<br>ATGACG                                                     |
| 250.       | L0266C03-3 | D930020B18Rik | RIKEN cDNA D930020B18 gene                                                                                      | L0266C03   | Mm.138048  | Chromosome 10                         |                                                                                                                                                                                                                                                                                                                      |
| 251.       | K0423B04-3 | Zfp91         | zinc finger protein 91                                                                                          | K0423B04   | Mm.212863  | Chromosome 19                         |                                                                                                                                                                                                                                                                                                                      |
| 252.       | J0403C04-3 |               | AU021859 Mouse unfertilized egg cDNA Mus musculus cDNA clone J0403C04 3', MRNA sequence                         | J0403C04   |            | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                      |
| 253.       | J0248E12-3 | 1700011I03Rik | RIKEN cDNA 1700011I03 gene                                                                                      | J0248E12   | Mm.78729   | No Chromosome location info available | CTGAAA.<br>GGAAAA<br>AAATAC.<br>TTAGGA/<br>AATATG/<br>GAAAAAC<br>GCGAGA<br>TGAAAA                                                                                                                                                                                                                                    |
| 254.       | J0908H04-3 | Rpl24         | ribosomal protein L24                                                                                           | J0908H04   | Mm.107869  | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                      |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL | GENE NAME                                                                                                                           | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------|------------|-------------|-------------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 255.       | K0205H10-3 | Madd        | MAP-kinase activating death domain                                                                                                  | K0205H10   | Mm.36410   | Chromosome 2                          | GAAAATG.<br>AATACAC/<br>TAGGACG/<br>AATATGG/<br>AGAAAGC/<br>GGACTGG.<br>GGAGGAG/<br>GTAAATA/<br>AGCTCCA/<br>ATTTATAC<br>ACAAAAA<br>TACCTATC<br>ACAGTGA<br>AAGAGAG<br>TGTTTAGT<br>TCAGGTT/<br>CTAAGGG<br>AAATGTT/<br>TAAAATG/<br>AAAGAAC<br>GAGGCAA<br>GGAGTGG                                                                                                                                                                                               |
| 256.       | C0507E09-3 | Gpr22       | G protein-coupled receptor 22                                                                                                       | C0507E09   | Mm.68486   | Chromosome 12                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 257.       | J0005B11-3 |             | Mus musculus transcribed sequence with weak similarity to protein ref:NP_083358.1 (M.musculus) RIKEN cDNA 5830411J07 [Mus musculus] | J0005B11   | Mm.249862  | Chromosome 7                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 258.       | L0201E08-3 | AW551705    | ESTs AW551705                                                                                                                       | L0201E08   | Mm.182670  | Chromosome 6                          | CCACATC/<br>GAAAGAA/<br>CACTTAT/<br>ATTGCCA/<br>ATAGGAC/<br>GAAAGTC<br>ATGAGAA/<br>CACACTT<br>ACGTGAA/<br>GGCGAGC<br>ACTGAAA/<br>GTCTATT/<br>CCTGTGA<br>AAAATGC<br>TGATCCA<br>CTAAATC<br>AACCTGC<br>GTAGATC<br>GCGAGA/<br>TGAAAA/<br>GAAAATC<br>AATACAC/<br>TAGGAC/<br>AATATGC<br>TGGAGG/<br>GATTGA/<br>CGATTGC<br>ATCGAA/<br>GAGAAA<br>ATGTTCA/<br>CTTCATC<br>TTTTCAC<br>ATAATA/<br>GAAAAG<br>GGTAAA/<br>ATCACTC<br>ACTGAA/<br>TGGAAA<br>GAAACA<br>TTGACG/ |
| 259.       | J0426H03-3 | AU023164    | ESTs AU023164                                                                                                                       | J0426H03   | Mm.221086  | Chromosome 4                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 260.       | C0649D06-3 | Cdkn2b      | cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)                                                                           | C0649D06   | Mm.269426  | Chromosome 4                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 261.       | J0421D03-3 | Rpl24       | ribosomal protein L24                                                                                                               | J0421D03   | Mm.107869  | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 262.       | K0643F07-3 |             | ESTs BQ563001                                                                                                                       | K0643F07   | Mm.25571   | Chromosome X                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 263.       | H3103C12-3 | Slamf1      | signaling lymphocytic activation molecule family member 1                                                                           | H3103C12   | Mm.103648  | Chromosome 1                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 264.       | J0416H11-3 | Pscdbp      | pleckstrin homology, Sec7 and coiled-coil domains,                                                                                  | J0416H11   | Mm.123225  | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                                                                                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 265.             | AF015770.1 | Rfng           | binding protein<br>radical fringe<br>gene homolog<br>(Drosophila)                                                                                                                                                                                                                          | AF015770      | Mm.871        | Chromosome 11            | AAAAATG<br>AAATCAC<br>CAAGCAC<br>CTGCAAA<br>CGGTGGA<br>GATAAGT<br>AGAATCT<br>GAAAA<br>TTTGAGA<br>AGGCATA<br>TGAAATA<br>GCAAAAA<br>ATACTGT<br>CGAGAC<br>GAAGAAA<br>AGGTGAA<br>CACTTTA<br>ACTTGGG<br>ACAGACC<br>ATATCCG<br>ATCATAA<br>TGTGGAA<br>ATATTGC<br>TTAAAAA<br>ACTATGG<br>GGAGAG<br>AAATGGC<br>AGAAAGC<br>AATGGCT<br>AAAATGC<br>AGTAGTC<br>AGAGGA<br>ATTTTAG<br>CTTTATT<br>CTTGACC<br>ATTTGAA<br>AAAAAG<br>GTCTGG |
| 266.             | C0933C05-3 |                | ESTs<br>BQ551952                                                                                                                                                                                                                                                                           | C0933C05      | Mm.217877     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 267.             | C0931A05-3 | E130304F04Rik  | RIKEN cDNA<br>E130304F04<br>gene                                                                                                                                                                                                                                                           | C0931A05      | Mm.38058      | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 268.             | J0030C02-3 | C77383         | ESTs C77383                                                                                                                                                                                                                                                                                | J0030C02      | Mm.43952      | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 269.             | H3061A07-3 | Srpk2          | serine/arginine-<br>rich protein<br>specific kinase<br>2                                                                                                                                                                                                                                   | H3061A07      | Mm.8709       | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 270.             | J0823B08-3 |                | AU041035<br>Mouse four-<br>cell-embryo<br>cDNA Mus<br>musculus<br>cDNA clone<br>J0823B08 3',<br>MRNA<br>sequence<br>Mus musculus<br>transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus] | J0823B08      |               | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 271.             | L0942H08-3 |                | Mus musculus<br>transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus]                                                                                                                     | L0942H08      | Mm.276728     | Chromosome 11            | GTGGAA<br>GAGATC<br>ACGTTT<br>TAGGAG<br>AATGAG<br>ATTAAAC                                                                                                                                                                                                                                                                                                                                                               |
| 272.             | C0280H06-3 | Mrpl50         | mitochondrial<br>ribosomal<br>protein L50                                                                                                                                                                                                                                                  | C0280H06      | Mm.30052      | Chromosome 4             | AAACCC<br>GTAGCC<br>GGCCCG<br>CACCAA<br>TTTTTA<br>AAGGA<br>ATTATG<br>TGTAAC<br>GAAGTA<br>CTGTGA<br>TCAACC<br>GATGA<br>AGAAGA<br>CTGAGC                                                                                                                                                                                                                                                                                  |
| 273.             | L0534E07-3 | 4632417D23     | hypothetical<br>protein<br>4632417D23                                                                                                                                                                                                                                                      | L0534E07      | Mm.105080     | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 274.             | U22339.1   | Il15ra         | interleukin 15<br>receptor, alpha                                                                                                                                                                                                                                                          | U22339        | 16169         | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME<br>chain                                                                                                      | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                     |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 275.             | L0533C12-3 |                | L0533C12-3<br>NIA Mouse<br>Newborn Heart<br>cDNA Library<br>Mus musculus<br>cDNA clone<br>L0533C12 3',<br>MRNA<br>sequence | L0533C12      |               | No Chromosome location<br>info available | AACCCCTT<br>ATAGGA<br>GACAAA<br>AACTCAC<br>CTGCCTT<br>TAAAAA<br>AGGCATC<br>AACCAA<br>GGCCAGC<br>AGTTAAC                                                                                                                                                                                                                                                                                             |
| 276.             | C0909E04-3 | Mvk            | mevalonate<br>kinase                                                                                                       | C0909E04      | Mm.28088      | Chromosome 5                             | ACAAGCC<br>GCCTCTG<br>CACCCGA<br>CCATCCT<br>AGAAGCC<br>AGTAT<br>CAAGTCA<br>AGAAGCC<br>CTTGGTG<br>AATTCTG<br>TTTGAAA<br>GGTCTTG<br>GGTCAAC<br>GTGCCAA<br>CTTTGTT<br>AAATCCT<br>CTGAATC<br>AGCCTG<br>AGTGGAA<br>TATAAGC<br>AACCCAG<br>GAGTCGC<br>ATTTCCA<br>ACTCAT<br>CTTCCCA<br>CCCACCG<br>TTGTCTA<br>TGCATGT<br>GTAAAAA<br>AAAAAAG<br>TGCCTGA<br>AAGAAAA<br>GCCAGAA<br>GAACCAT<br>ATCTTTA<br>TCTTCT |
| 277.             | J0093B09-3 | Bhmt2          | betaine-<br>homocysteine<br>methyltransfera<br>se 2                                                                        | J0093B09      | Mm.29981      | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 278.             | H3066D09-3 | BG068517       | ESTs<br>BG068517                                                                                                           | H3066D09      | Mm.250067     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 279.             | C0346F01-3 | BM197260       | ESTs<br>BM197260                                                                                                           | C0346F01      | Mm.222100     | Chromosome 9                             |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 280.             | K0125A06-3 | Hdac7a         | histone<br>deacetylase 7A                                                                                                  | K0125A06      | Mm.259829     | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 281.             | J0214H07-3 |                | C85807 Mouse<br>fertilized one-<br>cell-embryo<br>cDNA Mus<br>musculus<br>cDNA clone<br>J0214H07 3',<br>MRNA<br>sequence   | J0214H07      |               | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 282.             | C0309H10-3 | 5930412E23Rik  | RIKEN cDNA<br>5930412E23<br>gene                                                                                           | C0309H10      | Mm.45194      | No Chromosome location<br>info available | GTTAATA<br>TAACTGA<br>GCCCATA<br>CCGTGGT<br>GGTGTTG<br>CAGTG<br>GGAGGAC<br>ATCCTCA<br>CCTCATC<br>CCCAACA<br>ATAAAGT<br>TTTAAC<br>TCTGAAC<br>ACCCATC                                                                                                                                                                                                                                                 |
| 283.             | C0351C04-3 | 2610034E13Rik  | RIKEN cDNA<br>2610034E13<br>gene                                                                                           | C0351C04      | Mm.157778     | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                                     |
| 284.             | K0204G07-3 | Arf3           | ADP-<br>ribosylation                                                                                                       | K0204G07      | Mm.295706     | Chromosome 15: not<br>placed             |                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                     | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                    |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|----------------|------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | factor 3                                                                                                                                                         |               |                |                                          | ACCCCGT<br>TCAACA<br>TTCCAAA<br>TCTGG<br>AGGAGCC<br>TCCTTAT<br>TTGGAAT<br>TTCAGCC<br>ATCTCAC<br>TCTGT                                                                                                                                                              |
| 285.             | L0928B09-3 |                | transcribed<br>sequence with<br>strong<br>similarity to<br>protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | L0928B09      | Mm.217064      | Chromosome 10                            |                                                                                                                                                                                                                                                                    |
| 286.             | H3059A09-3 | C430004E15Rik  | RIKEN cDNA<br>C430004E15<br>gene                                                                                                                                 | H3059A09      | Mm.29587       | Chromosome 2                             | GAAAAA/<br>GAGATCT<br>CATGACA<br>GCCTGCA<br>ACATTTC<br>CCCTTCT<br>TTTGATT<br>CAGAAAC<br>CACCAAA<br>TGCCTTA<br>TATTTCT<br>AGGGGA<br>AGATACT<br>TACTGTC<br>AATGCAG<br>GACTCTA<br>AACAACC<br>AAAATG<br>AGAGAAC<br>CACTCCT<br>TCAAGAC<br>AGAGCAT<br>CAACCAA<br>GCTATTT |
| 287.             | C0949D03-3 |                | UNKNOWN<br>C0949D03                                                                                                                                              | C0949D03      | Data not found | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                    |
| 288.             | K0118A04-3 | Rgs1           | regulator of G-<br>protein<br>signaling 1                                                                                                                        | K0118A04      | Mm.103701      | Chromosome 1                             |                                                                                                                                                                                                                                                                    |
| 289.             | H3123F11-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus]           | H3123F11      | Mm.157781      | Chromosome 7                             |                                                                                                                                                                                                                                                                    |
| 290.             | H3154A06-3 | Gng13          | guanine<br>nucleotide<br>binding protein<br>13, gamma                                                                                                            | H3154A06      | Mm.218764      | Chromosome 17                            | TATGAGC<br>CCCACAC<br>TGTAAGG<br>ACTTTAT<br>AGACTTC<br>GGTGT<br>ATACCCC<br>CAACCTC<br>AAGAGGC<br>TAACCTG<br>GATAAAA<br>ATCAGG                                                                                                                                      |
| 291.             | L0534E01-3 |                | L0534E01-3<br>NIA Mouse<br>Newborn Heart<br>cDNA Library<br>Mus musculus<br>cDNA clone<br>L0534E01 3',<br>MRNA<br>sequence                                       | L0534E01      |                | Chromosome 9                             |                                                                                                                                                                                                                                                                    |
| 292.             | L0250B10-3 | Ap4m1          | adaptor-related<br>protein<br>complex AP-4,<br>mu 1                                                                                                              | L0250B10      | Mm.1994        | No Chromosome location<br>info available | TATCCTCC<br>AAAGATG<br>GGAGCCC<br>AGTGTTA<br>TAGAAGT<br>GTGAAA<br>TATTGTCC                                                                                                                                                                                         |
| 293.             | L0518G04-3 | BM123045       | ESTs                                                                                                                                                             | L0518G04      | Mm.221745      | Chromosome 3                             |                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                 |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | BM123045                                                                                                                                                           |               |               |                                          | GAAACCC<br>AACTACC<br>ATCTGGA<br>GAACAT1<br>TGCATT<br>TAAGGAC<br>GCCCTAC<br>CTACGAT<br>CTATCAC<br>AAAATTA<br>AAAGGG                                                                             |
| 294.             | J1020E03-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | J1020E03      | Mm.250157     | Chromosome 9                             |                                                                                                                                                                                                 |
| 295.             | X12616.1   | Fes            | feline sarcoma<br>oncogene                                                                                                                                         | X12616        | Mm.48757      | Chromosome 7                             | TCAAGGC<br>GTTTCTG<br>AAGCAAC<br>CCTGAAA<br>ACAACCA<br>AACATTG<br>GATTGCC<br>ACTTACA<br>ATAGAG1<br>AAGCCCA<br>AGCCTGA<br>GAGCCA<br>TTATTCC<br>GCCCCC<br>AGATGTT<br>CAACCGA<br>AGCGGTC<br>AAGAGC |
| 296.             | J0026H02-3 | C77164         | expressed<br>sequence<br>C77164                                                                                                                                    | J0026H02      | 97587         | Chromosome X                             |                                                                                                                                                                                                 |
| 297.             | H3154D11-5 | Taf7l          | TAF7-like<br>RNA<br>polymerase II,<br>TATA box<br>binding protein<br>(TBP)-<br>associated<br>factor                                                                | H3154D11      | Mm.103259     | Chromosome X                             | TTATTCC<br>GCCCCC<br>AGATGTT<br>CAACCGA<br>AGCGGTC<br>AAGAGC                                                                                                                                    |
| 298.             | H3054H04-3 | Kcnn4          | potassium<br>intermediate/sm<br>all conductance<br>calcium-<br>activated<br>channel,<br>subfamily N,<br>member 4                                                   | H3054H04      | Mm.9911       | Chromosome 7                             | AGCTCCA<br>AACTCAC<br>AGAACC/<br>TAAGTAC<br>GGACCG/<br>AAGGAC/                                                                                                                                  |
| 299.             | J0425B03-3 | R75183         | expressed<br>sequence<br>R75183                                                                                                                                    | J0425B03      | Mm.276293     | Chromosome 15                            | ACCATTA<br>TTTAAAA<br>CAAAAAA<br>CAGCAAC<br>GCCTTTG<br>GCCTCAA<br>CTTCATC<br>AACTCCA<br>AACTCCC<br>TAACCTG<br>CCAGCAC<br>CAGTT                                                                  |
| 300.             | C0930C02-3 | 0610037D15Rik  | RIKEN cDNA<br>0610037D15<br>gene                                                                                                                                   | C0930C02      | Mm.218714     | No Chromosome location<br>info available | CTGCACC<br>AGGAGCC<br>GTGAAGC<br>CAGCACT<br>CATGTTA<br>GAGTCT<br>CACTGG/<br>TGAACA1<br>TACAAGT                                                                                                  |
| 301.             | L0812A11-3 |                | ESTs BI793430                                                                                                                                                      | L0812A11      | Mm.261348     | No Chromosome location<br>info available |                                                                                                                                                                                                 |
| 302.             | J0243F04-3 | 9530020D24Rik  | RIKEN cDNA<br>9530020D24<br>gene                                                                                                                                   | J0243F04      | Mm.200585     | Chromosome 2                             |                                                                                                                                                                                                 |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 303.             | C0335A03-3 | I110035O14Rik  | RIKEN cDNA<br>I110035O14<br>gene                                                                                                                                                   | C0335A03      | Mm.202727     | Chromosome 12            | CACAGA/<br>CAGCAC/<br>TGTACT<br>ATAAGA/<br>TAGGAA/<br>ACTCCC/<br>AAAATA/<br>ACCTCA/<br>TGGGGA/<br>GCCCCA/<br>CTAAAT/<br>TACTTAT<br>ATTCTCT<br>TAGGACT<br>TCCTG<br>CAGTCA/<br>TCCAGA/<br>TACAAC/<br>GGAGAA/<br>AATGAC/<br>TCTCCT<br>CGTAGC/<br>GGTAGA/<br>CTGACC/<br>GCATAC/<br>TGGGTT/<br>AAGGAA/<br>GAGCCT/<br>CTACGA/<br>ATTTCAT<br>TTCAAG/<br>TTTIGAC<br>TCAAG<br>GAGTCT/<br>TATTGCG<br>ACAAGC/<br>GCCCAA/<br>ATTTCAA<br>AAGAAA/<br>AGCATC/<br>AAAGCA/<br>AACTCG/<br>AAGCAA/<br>TGTCTT/<br>GTCAAA/<br>GGGAAA/<br>AGCAAA/<br>CAAACCT/<br>ACCACA/<br>CCTGTTA<br>TGGTGGC |
| 304.             | H3003B10-3 | BG063111       | ESTs<br>BG063111                                                                                                                                                                   | H3003B10      | Mm.100527     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 305.             | U97073.1   | Prtn3          | proteinase 3                                                                                                                                                                       | U97073        | Mm.2364       | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 306.             | K0300D08-3 | Afmid          | arylformamidase                                                                                                                                                                    | K0300D08      | Mm.169672     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 307.             | H3029H06-3 | Sf3b2          | splicing factor<br>3b, subunit 2                                                                                                                                                   | H3029H06      | Mm.196532     | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 308.             | H3074D09-3 | Drg2           | developmentally regulated<br>GTP binding<br>protein 2                                                                                                                              | H3074D09      | Mm.41803      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 309.             | K0647G12-3 | Plek           | pleckstrin                                                                                                                                                                         | K0647G12      | Mm.98232      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 310.             | H3137A08-3 |                | Mus musculus<br>transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>pir:S12207<br>(M.musculus)<br>S12207<br>hypothetical<br>protein (B2<br>element) -<br>mouse | H3137A08      | Mm.197271     | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 311.             | C0166D06-3 | Slc38a3        | solute carrier<br>family 38,<br>member 3                                                                                                                                           | C0166D06      | Mm.30058      | Chromosome 9             | ACACAG/<br>AGAAAA/<br>GGCCTG/<br>CATCCCC<br>CCTGCTC<br>ACCACA/<br>CGACCA/<br>CCTGGG/                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 312.             | K0406B07-3 | Sirt7          | sirtuin 7 (silent<br>mating type                                                                                                                                                   | K0406B07      | Mm.259849     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                             | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUEN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------|------------|---------------|-----------------------------------------------------------------------|------------|------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 313.       | H3085D10-3 | Gda           | information regulation 2, homolog 7 (S. cerevisiae) guanine deaminase | H3085D10   | Mm.45054   | Chromosome 19                         | AACACCCG<br>GAACGGG<br>CAGAAAC<br>TGAGTGA<br>GTTTAGG<br>TTTCCATI<br>TCTTATA/<br>AGAAACC<br>AGGCAGT<br>AGTTC<br>TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA |
| 314.       | H3099C09-3 | Igf1          | insulin-like growth factor 1                                          | H3099C09   | Mm.268521  | Chromosome 10                         | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 315.       | H3099B07-5 | 2610028H24Rik | RIKEN cDNA 2610028H24 gene                                            | H3099B07   | 76964      | No Chromosome location info available | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 316.       | H3114H10-3 | Rec8L1        | REC8-like 1 (yeast)                                                   | H3114H10   | Mm.23149   | Chromosome 14                         | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 317.       | L0703E03-3 | Lipc          | lipase, hepatic                                                       | L0703E03   | Mm.362     | Chromosome 9                          | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 318.       | H3074H08-3 | BG069302      | ESTs BG069302                                                         | H3074H08   | Mm.11484   | Chromosome 7: not placed              | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 319.       | K0443D01-3 | Baz1b         | bromodomain adjacent to zinc finger domain, 1B                        | K0443D01   | Mm.40331   | Chromosome 5                          | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 320.       | J0409E10-3 | AU022163      | ESTs AU022163                                                         | J0409E10   | Mm.188475  | Chromosome 16                         | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 321.       | L0528E01-3 | BM123655      | EST BM123655                                                          | L0528E01   | Mm.216782  | Chromosome 9                          | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 322.       | L0031B11-3 | Alcam         | activated leukocyte cell adhesion molecule                            | L0031B11   | Mm.2877    | Chromosome 16                         | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |
| 323.       | G0115A06-3 | Fem1a         | feminization 1 homolog a (C.                                          | G0115A06   | Mm.27723   | Chromosome 17                         | TCGAAAC<br>ACCAAA<br>ATAATAA<br>AATAACA<br>AAAGATC<br>ATTTC<br>TGCTACC<br>AGGACCA<br>ATGGATC<br>ACGGAGT<br>AGAGCTC<br>AGCAGA/<br>CGGAGCT<br>AGAACCC<br>CTCTCTC<br>TGGCTAC<br>AGAACTC<br>GTTTAT<br>ATAAAG<br>TTCCCA<br>CTGGGC<br>GAATTAC<br>AATAAA<br>TTCCCTC<br>ACTTTC/<br>TGAATC/<br>AGCCTG<br>AGATCT<br>AGAAAC<br>CCCCAA<br>GACAA<br>AGGGAC<br>AAAAAC<br>GGAAG/<br>AAAATC<br>TTTTTT/<br>GCCCAA<br>AGAAA/<br>CTCTAT<br>GAGAT/<br>TATTAA<br>TAGTAC<br>CTCCAC<br>AAGTC<br>AATAGC<br>GATTAC<br>TCTCGC<br>TGCTC/<br>TTTCTG<br>CCACTC<br>CCATT<br>CAGAT<br>GTATA<br>ACTGG<br>ATACA/<br>GCTGA                                                                                                                |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME<br><br>(elegans)                                             | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------|------------|----------------|---------------------------------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 324.             | L0947C07-3 | Mal            | myelin and<br>lymphocyte<br>protein, T-cell<br>differentiation<br>protein | L0947C07      | Mm.39040      | Chromosome 2             | TTGAGTC<br>ATGAGG<br>AATAAG<br>CCAGCA<br>TCTTATA<br>CAACAA<br>GAACCC<br>TTACAC<br>AGCAGC<br>ACGAGT<br>CTGAAT<br>CACACC<br>GAGACT<br>TGAGCG<br>CCAAAT<br>TGAAT<br>GTTCTC<br>GAGTGC<br>AACCCA<br>GTCTGA<br>TGAAGG<br>AACTGT<br>AAACAC<br>ACTTGA<br>CCATGA<br>CTCAAA<br>TTCTATC<br>TTTGGA<br>TCATGG<br>TAACTA<br>ATAAAG<br>ACACCC<br>GAAGCA<br>GCGTCC<br>GGACAC<br>ACACTG<br>ACAGAG<br>CAACTT<br>TTTGTG<br>CAGCAA<br>AGGAAA<br>GGGGTT<br>CTCTCA<br>TTAAAG<br>GCCTAA<br>AGGTGT<br>CTCAAG<br>GCCAAC<br>CCGTTT<br>ACCCTG<br>TGATCG<br>TTCAT<br>TCTGTAC<br>CCGAAA<br>GAGTCC<br>ATTCTT<br>TATCCA<br>CTCTGA<br>TTCTGG<br>TATTTC<br>TCTTTA<br>AGTTCA<br>AGTGTG<br>AAGAA<br>GCAGAT<br>AACTAG |
| 325.             | H3101A05-3 | AU040576       | expressed<br>sequence<br>AU040576                                         | H3101A05      | Mm.26700      | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 326.             | H3064E10-3 | BG068353       | ESTs<br>BG068353                                                          | H3064E10      | Mm.35046      | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 327.             | K0505H05-3 | Ian6           | immune<br>associated<br>nucleotide 6                                      | K0505H05      | Mm.24781      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 328.             | H3082E12-3 | Ptpre          | protein tyrosine<br>phosphatase,<br>receptor type, E                      | H3082E12      | Mm.945        | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 329.             | H3088A06-3 | 2310047N01Rik  | RIKEN cDNA<br>2310047N01<br>gene                                          | H3088A06      | Mm.31482      | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 330.             | K0635B07-3 | Ccr5           | chemokine (C-<br>C motif)<br>receptor 5                                   | K0635B07      | Mm.14302      | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 331.             | C0153A12-3 | 1110025F24Rik  | RIKEN cDNA<br>1110025F24<br>gene                                          | C0153A12      | Mm.28451      | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 332.             | C0143E02-3 | BC022145       | cDNA sequence<br>BC022145                                                 | C0143E02      | Mm.200891     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 333.             | L0863F12-3 | Nr2c2          | nuclear receptor<br>subfamily 2,<br>group C,<br>member 2                  | L0863F12      | Mm.193835     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 334.             | H3045F02-3 | LOC214424      | hypothetical<br>protein                                                   | H3045F02      | Mm.31129      | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                                                                             | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                            |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | LOC214424                                                                                                                                                                                                                                                |               |               |                          | CTGTCAT<br>TTCTAAA<br>ACCAACT<br>ATTAC<br>CTTAAAA<br>GAGATAC<br>TTACTCT<br>CAGCAA<br>GTTAAGA<br>AGAATG<br>CTTCCTG<br>ATTACCA<br>GAAAACC<br>ATGGCCC<br>CCATATA<br>GAAGTT<br>GTAACGC<br>CTGGGGG<br>AGGTTAT<br>ACATATA<br>CAAACGT<br>CAAGAG<br>TCCCCAC<br>CATGGGG<br>TCAAGAA<br>CACCATT<br>GAAAGGT<br>AAAAA<br>GCGCAGA<br>AAACCAA<br>GGAGCCA<br>ATTGGTG<br>CAACCTA<br>CACCTTC |
| 335.             | H3035G05-3 | BG065832       | ESTs<br>BG065832                                                                                                                                                                                                                                         | H3035G05      | Mm.154695     | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                            |
| 336.             | H3137D02-3 | Hnrpl          | heterogeneous<br>nuclear<br>ribonucleoprote<br>in L                                                                                                                                                                                                      | H3137D02      | Mm.9043       | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                            |
| 337.             | H3097F07-3 | AU040829       | expressed<br>sequence<br>AU040829                                                                                                                                                                                                                        | H3097F07      | Mm.134338     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                            |
| 338.             | J0029C02-3 | Frag1-pending  | FGF receptor<br>activating<br>protein 1                                                                                                                                                                                                                  | J0029C02      | Mm.259795     | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                            |
| 339.             | BB416014.1 |                | Mus musculus<br>B6-derived<br>CD11 +ve<br>dendritic cells<br>cDNA, RIKEN<br>full-length<br>enriched<br>library,<br>clone:F730035<br>A01<br>product:similar<br>to SWI/SNF<br>COMPLEX<br>170 KDA<br>SUBUNIT<br>[Homo<br>sapiens], full<br>insert sequence. | BB416014      | Mm.24449      | Chromosome 10            | annexin A4                                                                                                                                                                                                                                                                                                                                                                 |
| 340.             | H3087E01-3 | Anxa4          |                                                                                                                                                                                                                                                          | H3087E01      | Mm.259702     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                            |
| 341.             | H3088E08-3 | BG070548       | ESTs<br>BG070548                                                                                                                                                                                                                                         | H3088E08      | Mm.11161      | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                            |
| 342.             | AF179424.1 |                | Mus musculus<br>13 days embryo<br>male testis<br>cDNA, RIKEN<br>full-length<br>enriched<br>library,<br>clone:6030408                                                                                                                                     | AF179424      | Mm.1428       | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                            |

| SEQ ID NO: | CLONE ID    | GENE SYMBOL   | GENE NAME                                                                                           | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------|-------------|---------------|-----------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 343.       | J0258C01-3  |               | M17 product:GATA binding protein 4, full insert sequence<br>Mus musculus mRNA for mKIAA1335 protein | J0258C01   | Mm.275718  | Chromosome 2                          | CCCTAGT<br>TTTCTGA<br>TCAGAAC<br>AATAACT<br>GTAGTCC<br>GCTTT<br>GTAGCCA<br>GCCACAA<br>ACAAATG<br>CTGTGAA<br>ATATGGA<br>TTTATT<br>GGCTCCA<br>TGAAGCT<br>TTAAGCT<br>AGATTTT<br>AAACGCT<br>AAAGC<br>TGCTGGG<br>TAGAACC<br>GACATAG<br>ATGGATA<br>GCAACCG<br>TGGCAAA<br>AACGCAA<br>GCAAGAA<br>AACAAAG<br>GGAACAA<br>GCAGAA<br>TCCCGCC<br>TGTTTTCT<br>GACCAAA<br>ATGACAA<br>GCAGAAA<br>GAACTGA<br>AATTGAT<br>CCCACCA<br>ATATAGA<br>ACTGTGA<br>GACCATA<br>GGTCCTG<br>TTTAAT<br>GTATGAC<br>AACCAGA<br>AGGCTCT<br>GCTGAAC<br>TAACCGG<br>AAAACG<br>CTTCTGG<br>CTTACAT<br>GACTGAT<br>GAAACCA<br>CATTCTT<br>TTTGAA<br>GCAGGGT<br>ACTTTCT<br>GCCTGAA<br>CTTCATT<br>TTGGCACT<br>TAACA |
| 344.       | K0507B09-3  |               | ESTs<br>BM238095                                                                                    | K0507B09   | Mm.218038  | Chromosome 9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 345.       | L0846F07-3  | BM117131      | ESTs<br>BM117131                                                                                    | L0846F07   | Mm.216977  | Chromosome 9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 346.       | U48866.1    | CEBPE         | CCAAT/enhancer binding protein (C/EBP), epsilon                                                     | U48866     | Hs.158323  | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 347.       | K0301B06-3  | Fech          | ferrochelatase                                                                                      | K0301B06   | Mm.217130  | Chromosome 18                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 348.       | NM_009756.1 | Bmp10         | bone morphogenetic protein 10                                                                       | NM_009756  | Mm.57171   | Chromosome 6                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 349.       | NM_010100.1 | Edar          | ectodysplasin-A receptor                                                                            | NM_010100  | Mm.174523  | Chromosome 10                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 350.       | G0115E06-3  | C430014D17Rik | RIKEN cDNA C430014D17 gene                                                                          | G0115E06   | Mm.103389  | Chromosome 3                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 351.       | L0266D11-3  | Ppp3ca        | protein phosphatase 3, catalytic subunit, alpha isoform                                             | L0266D11   | Mm.80565   | Chromosome 3                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 352.       | L0526F10-3  |               | Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone:A830020         | L0526F10   | Mm.215689  | Chromosome X                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| SEQ ID NO: | CLONE ID    | GENE SYMBOL   | GENE NAME                                                                                                                 | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUENCE                                                            |
|------------|-------------|---------------|---------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|---------------------------------------------------------------------------|
| 353.       | H3047C10-3  | Slc6a6        | C21 product: unknown EST, full insert sequence. solute carrier family 6 (neurotransmitter transporter, taurine), member 6 | H3047C10   | Mm.200518  | Chromosome 6                          | TTAGCAC<br>GAAAAGC<br>GAACGTC<br>TTGCCTC<br>AGAAAATA<br>TGGCTC            |
| 354.       | K0322G06-3  | BC042620      | cDNA sequence BC042620                                                                                                    | K0322G06   | Mm.152289  | Chromosome 17                         | ACACAGC<br>ACAACTA<br>TGGGACA<br>TATCTGG<br>AAGAGAC<br>ACTAAT             |
| 355.       | NM_009580.1 | Zp1           | zona pellucida glycoprotein 1                                                                                             | NM_009580  | Mm.24767   | Chromosome 19                         | CAATGGC<br>TCTGTCA<br>GGTGTCC<br>AAGGGTG<br>ACTACAG                       |
| 356.       | H3150E08-3  | Map4k5        | mitogen-activated protein kinase kinase kinase 5                                                                          | H3150E08   | Mm.260244  | Chromosome 12                         | ACAAGTA<br>AAAGTAG<br>ACACAGT<br>GGGATAA<br>ATCTGGA<br>TGATCAG            |
| 357.       | J0059G03-3  | C79059        | ESTs C79059                                                                                                               | J0059G03   | Mm.249888  | Chromosome 4                          | GAGTTA<br>CACCTGG<br>ACAGCTA<br>GATTCTA<br>GACAGGG<br>AGCATCT             |
| 358.       | U93191.1    | Hdac2         | histone deacetylase 2                                                                                                     | U93191     | 15182      | Chromosome 10                         | CAAAGT<br>TATTAAAC<br>GGAGATA<br>GGAGTCT<br>TTAACCTC<br>GTAACCTC          |
| 359.       | H3033C04-5  |               | H3033C04-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3033C04 5', MRNA sequence                                | H3033C04   |            | No Chromosome location info available | GTAGTT<br>TTCCTCCC<br>ATGGAGT<br>TCTTCAAA<br>CAGCTCCC<br>AGATCTA<br>GATAT |
| 360.       | H3085C01-3  | 2700038N03Rik | RIKEN cDNA 2700038N03 gene                                                                                                | H3085C01   | Mm.21836   | Chromosome 5                          | TATGTCTT<br>ACTGGACC<br>ACTACTGC<br>ACTCCAAA<br>ACCGTTGT<br>CTACAA        |
| 361.       | J0412G02-3  | BB336629      | ESTs BB336629                                                                                                             | J0412G02   | Mm.208743  | Chromosome 11                         | AGTAAAGC<br>ACCGGAA<br>TAAATCCT<br>TTAGGATA<br>AAGGAAT<br>GGGATGG         |
| 362.       | K0527H09-3  | BM239048      | ESTs BM239048                                                                                                             | K0527H09   | Mm.217288  | Chromosome 11                         | GAATGTCT<br>ACATGACC<br>CAGTTAGG<br>CACTGAAC<br>AGGAGTAC                  |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                                                                               | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                                                   |
|------------|------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 363.       | H3009C10-3 | Serpib9b      | serine (or cysteine) proteinase inhibitor, clade B, member 9b                                                                                                                           | H3009C10   | Mm.45371   | Chromosome 13                         | AAACTC<br>GCTTCTA<br>CTCTTGT<br>ATATGTG<br>TATCCAG<br>AGGATTT<br>AAGCA                                                                                                                                                                                                                                                                                                                        |
| 364.       | H3142D11-3 |               | Mus musculus mRNA similar to hypothetical protein FLJ20811 (cDNA clone MGC:27863 IMAGE:34925 16), complete cds                                                                          | H3142D11   | Mm.113272  | Chromosome X                          | CTGTCTA<br>CTGAACC<br>AGCAGAA<br>ACACCCA<br>AGAGCTT<br>CAAATA                                                                                                                                                                                                                                                                                                                                 |
| 365.       | H3094B07-3 |               | Mus musculus transcribed sequence with weak similarity to protein sp:P11369 (M.musculus) POL2_MOUSE Retrovirus-related POL polyprotein [Contains: Reverse transcriptase ; Endonuclease] | H3094B07   | Mm.173357  | Chromosome 14                         | AAAGGAC<br>GCATCAG<br>TCTGATA<br>GCTGAGG<br>AGATTGA<br>TGGGATT                                                                                                                                                                                                                                                                                                                                |
| 366.       | J0068F09-3 | C79588        | ESTs C79588                                                                                                                                                                             | J0068F09   | Mm.234023  | No Chromosome location info available | TGACTGG<br>ACCACCC<br>CTGAGTT<br>ATCTCAC<br>GGAACCT<br>GTTTCC<br>GGATCAG<br>ATGCACC<br>CTTTCCAT<br>TACATTTA<br>TCTTTTAC<br>TCAACC<br>TTGAGAC<br>AAGAAAT<br>AAACTCA<br>AAGATTAC<br>CCAGTGT<br>GTCATGG<br>GTCTCCTT<br>GTTATTGC<br>CCCAACAC<br>TAAGTCCC<br>TCAACAGC<br>TTCTA<br>CACAGCTC<br>GTAGTCAT<br>TCCAGTGA<br>GTAAGAAC<br>TTTATGT<br>TCTCTA<br>AACTTAAA<br>TCTCCAC<br>CTACCCCA<br>GATACTGC |
| 367.       | H3039B03-5 | E030024M05Rik | RIKEN cDNA E030024M05 gene                                                                                                                                                              | H3039B03   | Mm.5675    | Chromosome 12                         |                                                                                                                                                                                                                                                                                                                                                                                               |
| 368.       | H3068B03-3 | BG068673      | ESTs BG068673                                                                                                                                                                           | H3068B03   | Mm.11978   | Chromosome 1                          |                                                                                                                                                                                                                                                                                                                                                                                               |
| 369.       | C0250F05-3 | BM203195      | ESTs BM203195                                                                                                                                                                           | C0250F05   | Mm.228379  | Chromosome 12                         |                                                                                                                                                                                                                                                                                                                                                                                               |
| 370.       | H3110C11-3 | MIph          | melanophilin                                                                                                                                                                            | H3110C11   | Mm.17675   | Chromosome 1                          |                                                                                                                                                                                                                                                                                                                                                                                               |
| 371.       | H3121F01-3 | Wnt4          | wingless-related MMTV integration site 4                                                                                                                                                | H3121F01   | Mm.20355   | Chromosome 4                          |                                                                                                                                                                                                                                                                                                                                                                                               |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                               | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A] | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                |
|------------|------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------------|------------|------------|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 372.       | J1012G09-3 | Brd3          | bromodomain containing 3                                                                                                                | J1012G09   | Mm.28721   | Chromosome 2          | TATTTTTC<br>TGGT<br>CAGCAGAG<br>GGCTCCC<br>AGAAGGC<br>CAGCACA<br>ACAGCCA<br>GGATGTG<br>GGCTTCA<br>TAAGTGG<br>CTATTTT<br>TATTTAC<br>ATATGGT<br>AAATAA<br>CGCTCAG<br>AGAAAAGC<br>AAGGACA<br>ACTTGAT<br>CAAAGTC<br>CCAGTTA<br>GTCTGAA<br>CTATTAT<br>CCATCCA<br>CAACTGA<br>AGGGAGA<br>CTTTTG<br>AAAGAAT<br>AGAACGA<br>ATAGGTG<br>TGTAGTT<br>TACAGAA<br>AGATGCC |
| 373.       | L0952B09-3 | Usp49         | ubiquitin specific protease 49                                                                                                          | L0952B09   | Mm.25072   | Chromosome 17         |                                                                                                                                                                                                                                                                                                                                                            |
| 374.       | K0131B12-3 | Il4ra         | interleukin 4 receptor, alpha                                                                                                           | K0131B12   | Mm.233802  | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                            |
| 375.       | H3046E09-3 | Nfatc2ip      | nuclear factor of activated T-cells, cytoplasmic 2 interacting protein                                                                  | H3046E09   | Mm.1389    | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                            |
| 376.       | K0520B05-3 |               | transcribed sequence with weak similarity to protein pir:I58401 (M.musculus) I58401 protein-tyrosine kinase (EC 2.7.1.112) JAK3 - mouse | K0520B05   | Mm.221547  | Chromosome 14         |                                                                                                                                                                                                                                                                                                                                                            |
| 377.       | K0315G05-3 | Stat5a        | signal transducer and activator of transcription 5A                                                                                     | K0315G05   | Mm.4697    | Chromosome 11         | AAACCAC<br>AGTGTGA<br>GCCACG<br>TTGTAGT<br>CTGTTCA<br>AACAAT<br>GCACTCC<br>TGATTCT<br>GACTTTG<br>ACACATA<br>AAGTACT<br>ATTTG<br>ACTGTAT<br>TCCATGT<br>CTGACCA<br>AAGGCAA<br>GTATCAA<br>GAGAAA<br>GTGTTTG<br>AAAACCC<br>CCTCGGA<br>TTTAAAG<br>GGTTTTG<br>GTTGT                                                                                              |
| 378.       | H3086F07-3 | BC003332      | cDNA sequence BC003332                                                                                                                  | H3086F07   | Mm.100116  | Chromosome 6          |                                                                                                                                                                                                                                                                                                                                                            |
| 379.       | H3156A10-5 | Ctsd          | cathepsin D                                                                                                                             | H3156A10   | Mm.231395  | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                            |
| 380.       | C0890D02-3 |               | C0890D02-3 NIA Mouse Blastocyst cDNA Library (Long) Mus musculus cDNA clone C0890D02 3', MRNA sequence                                  | C0890D02   |            | Chromosome 18         |                                                                                                                                                                                                                                                                                                                                                            |
| 381.       | L0245G03-3 | 6430519N07Rik | RIKEN cDNA 6430519N07                                                                                                                   | L0245G03   | Mm.149642  | Chromosome 6          | CTCTCGAC<br>ATATAAA                                                                                                                                                                                                                                                                                                                                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME<br>gene                                                                                                                                                                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                               |
|------------------|------------|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 382.             | J0447A10-3 |                | Mus musculus<br>cDNA clone<br>IMAGE:12820<br>81, partial cds                                                                                                                                                            | J0447A10      | Mm.202311     | Chromosome 11                            | CAGTACC<br>TAAAGAGC<br>ATAAGTC<br>GCAAAGC<br>TATGGTA<br>TTTAGGG<br>GTCAGTT<br>ATGGGG/<br>ATTTGT/<br>AAACC<br>CTGGCTC<br>TGGCAAC<br>CATACTT<br>TTTAATA/<br>GAAACA/<br>ATTCATA                                                                                                                                                                    |
| 383.             | J1031A09-3 |                | Mus musculus<br>transcribed<br>sequence with<br>weak similarity<br>to protein<br>pir:I58401<br>(M.musculus)<br>I58401 protein-<br>tyrosine kinase<br>(EC 2.7.1.112)<br>JAK3 - mouse<br>RIKEN cDNA<br>A630084M22<br>gene | J1031A09      | Mm.235234     | No Chromosome location<br>info available | TTTGACC<br>GAAATAC<br>TTCATCT/<br>CAACACA<br>CCAGTAA<br>CACTG<br>CCTGTT/<br>TATCCTG/<br>CCACATA<br>CAAAGTT<br>ATACTAA<br>GAGAT                                                                                                                                                                                                                  |
| 384.             | L0072H04-3 | A630084M22Rik  |                                                                                                                                                                                                                         | L0072H04      | Mm.27968      | Chromosome 1                             | CTGGAAC<br>CACTGCC/<br>ACACTTG/<br>GAAATGC<br>GTTTGCC/<br>TTAAGT<br>CCTGGAG<br>CCACCTG.<br>TCCCTGA/<br>GGGTCAG<br>GCCTTGG<br>GGCCA<br>AAATGAG<br>CAGATTA/<br>AATTACC.<br>CCACCAA<br>ACCCCTC/<br>TCCTTG<br>CAGATAG<br>ACAGCAG<br>ATTTTCTT<br>TCCTGAA/<br>AATACCA/<br>CTCAAC<br>GGTGCCA.<br>CGGCCAT/<br>CTGAACA.<br>ATCGTCA/<br>GGAAGAA<br>TTGACC |
| 385.             | J0050E03-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_081764.<br>1 (M.musculus)<br>RIKEN cDNA<br>5730493B19<br>[Mus musculus]<br>TYRO3 protein<br>tyrosine kinase<br>3                                | J0050E03      | Mm.37806      | Chromosome 14                            |                                                                                                                                                                                                                                                                                                                                                 |
| 386.             | H3039C11-3 | Tyro3          |                                                                                                                                                                                                                         | H3039C11      | Mm.2901       | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                 |
| 387.             | C0324F11-3 | 6720458F09Rik  | RIKEN cDNA<br>6720458F09<br>gene                                                                                                                                                                                        | C0324F11      |               | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                 |
| 388.             | L0018F11-3 | AW547199       | ESTs<br>AW547199                                                                                                                                                                                                        | L0018F11      | Mm.182611     | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                 |
| 389.             | X69902.1   | Itga6          | integrin alpha 6                                                                                                                                                                                                        | X69902        | Mm.225096     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                 |
| 390.             | H3105A09-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_416488.<br>1 (E. coli)                                                                                                                          | H3105A09      | Mm.174047     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                 |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                       | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------|---------------|----------------|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 391.             | H3159F01-5 |                | putative<br>transport<br>protein,<br>shikimate<br>[Escherichia<br>coli K12]<br>UNKNOWN<br>H3159F01 | H3159F01      | Data not found | No Chromosome location<br>info available | CCAAAAC<br>GCCAACA<br>CGACAAC<br>CCCACAG<br>ACCCGGA<br>AAACCCA<br>TTTCAAC<br>CCATTAT<br>AGATTTA<br>CATCATT<br>AAACATG<br>CCAGAG<br>TGGAGAC<br>GTTGAC<br>CCATCTA<br>ACTGGCG<br>CAAGAGA<br>TGAAGTT<br>GATACAA<br>CATCTGT<br>CAAGGAG<br>TCATTTG<br>ACAAAAC<br>CAAGAGA<br>AACTAGA<br>CATAGAT<br>AGGACTC<br>CCATGAT<br>ACACTGG<br>ATGTTCT<br>ATCTCAA<br>TCTATCC<br>GGAAACA<br>TGAATCA<br>CACGACT<br>TGTGTG<br>AGAGGAG<br>CACTTGA<br>AATTAAA<br>TAAACAT<br>CCACTAA<br>TTTTAT<br>CTGCCGC<br>ACAAAGG<br>CTGAACC<br>TCATATTC<br>TAAATCA<br>GAGTTT<br>AAGCTGA<br>AAACGGC<br>ACAATAC<br>ATAGATA<br>CAACCGA<br>CTCAAGG<br>GACTTGGC<br>AACAATGC<br>CTCCATA<br>CAAAACTC<br>TTCCATGC<br>ACTTGCT |
| 392.             | K0522B04-3 | F5             | coagulation<br>factor V                                                                            | K0522B04      | Mm.12900       | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 393.             | C0123F08-3 | AI843918       | expressed<br>sequence<br>AI843918                                                                  | C0123F08      | Mm.143742      | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 394.             | H3067G08-3 | BG068642       | ESTs<br>BG068642                                                                                   | H3067G08      | Mm.250079      | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 395.             | K0349B03-3 | Stam2          | signal<br>transducing<br>adaptor<br>molecule (SH3<br>domain and<br>ITAM motif) 2                   | K0349B03      | Mm.45048       | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 396.             | C0620D11-3 | Bid            | BH3 interacting<br>domain death<br>agonist                                                         | C0620D11      | Mm.34384       | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 397.             | C0189H10-3 | 4930486L24Rik  | RIKEN cDNA<br>4930486L24<br>gene                                                                   | C0189H10      | Mm.19839       | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 398.             | H3140A02-3 | Slc9a1         | solute carrier<br>family 9<br>(sodium/hydrog<br>en exchanger),<br>member 1                         | H3140A02      | Mm.4312        | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 399.             | K0645B04-3 | Smc411         | SMC4<br>structural<br>maintenance of<br>chromosomes<br>4-like 1 (yeast)                            | K0645B04      | Mm.206841      | Chromosome 3                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 400.             | C0300G08-3 | 6720460I06Rik  | RIKEN cDNA<br>6720460I06<br>gene                                                                   | C0300G08      | Mm.28865       | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

| SEQ ID NO: | CLONE ID    | GENE SYMBOL   | GENE NAME                                                                                                                         | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A] | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------|-------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------|------------|------------|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 401.       | M59378.1    | Tnfrsf1b      | tumor necrosis factor receptor superfamily, member 1b                                                                             | M59378     | Mm.2666    | Chromosome 4          | AGCAGGC<br>AATTTGA<br>TGACCTA<br>ACATTCC<br>GGATGGC<br>TCCAGAA<br>AGCTCCA<br>AACAGAT<br>ACACAGC<br>TGGGAAC<br>CTGGGGA<br>CCATGAA<br>ACTAGCT<br>TGTAAG<br>CAAATCG<br>CTGAGTC<br>CACATAT<br>ACGGACA<br>GTAGGGT<br>ATACACC<br>CTACCGC<br>ATGAACC<br>AATTTTG<br>AGACA<br>TCCCCAC<br>AATTATC<br>CTAGTGG<br>AGGCCAC<br>ACAGGTT<br>TTGTT<br>TATGTGC<br>GCTGGAG<br>GGTTATA<br>GTACACT<br>GGCCAAT<br>TAGGA<br>CCACACT<br>GGAGACA<br>TCTGCCA<br>TGCATCA<br>TCAAACC<br>ACTTCT |
| 402.       | NM_009399.1 | Tnfrsf11a     | tumor necrosis factor receptor superfamily, member 11a                                                                            | NM_009399  | Mm.6251    | Chromosome 1          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 403.       | C0168E12-3  | 2810442I22Rik | RIKEN cDNA 2810442I22 gene                                                                                                        | C0168E12   | Mm.103450  | Chromosome 10         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 404.       | L0228H10-3  | C1r           | complement component 1, r subcomponent                                                                                            | L0228H10   | Mm.24276   | Chromosome 6          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 405.       | H3088B10-3  | BG070515      | ESTs BG070515                                                                                                                     | H3088B10   | Mm.11092   | Chromosome 11         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 406.       | K0409D10-3  | Lrrc5         | leucine-rich repeat-containing 5                                                                                                  | K0409D10   | Mm.23837   | Chromosome 5          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 407.       | H3056D02-3  |               | transcribed sequence with moderate similarity to protein ref:NP_079108.1 (H.sapiens) hypothetical protein FLJ22439 [Homo sapiens] | H3056D02   | Mm.9706    | Chromosome 12         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 408.       | J0430F08-3  | AU023357      | ESTs AU023357                                                                                                                     | J0430F08   | Mm.173615  | Chromosome 6          | TCGGTTGA<br>GATTCCAC<br>GGAGAAG<br>ATCAAGG<br>AGTAAAC<br>AGAGCAT<br>GAGTGCT<br>TGGTTGTT<br>GACCGTA<br>ATAGTCCT<br>TCAGACAC<br>GATTCTA<br>AACTGTCA<br>AATCCAAC<br>CCTTCATG<br>AAAGTTCC<br>GTCAGTAC<br>TAGAA                                                                                                                                                                                                                                                         |
| 409.       | H3158C06-3  | 2810457I06Rik | RIKEN cDNA 2810457I06 gene                                                                                                        | H3158C06   | Mm.133615  | Chromosome 9          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 410.       | M85078.1    | Csf2ra        | colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)                                                | M85078     | Mm.255931  | Chromosome 19         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 411.             | C0145E06-3 | Satb1          | special AT-rich<br>sequence<br>binding protein<br>1                                                                                                                                | C0145E06      | Mm.289605     | Chromosome 5             | ACTCTCA<br>TAAAGCC<br>CATCTCA<br>TCCTTGC<br>ACACAC<br>TAGGA<br>CAGACTG<br>GGAAATT<br>AGAAAAC<br>AACCTTT<br>CTATGAA<br>ATGGCTG<br>CTGAGAA<br>CTACTAC<br>TCTCTTT<br>ACCAACA<br>AGTGCCA<br>GTGGTT<br>AGCGACA<br>AACCAAA<br>ACTCAAA<br>AAAATCC<br>AAAACCTG<br>GTGAGGG<br>GTTTGTA<br>TAAAAGA<br>CCAGTGA<br>ATCCTAT<br>TTCTGGG<br>AATGA<br>ACTTAGA<br>AACAGCA<br>AGCATCA<br>CTTAAGT<br>AAGCAAA<br>CTAGTC<br>TAAACCA<br>TAAACTG<br>CTCCAGT<br>TTAGAAT<br>TGAAGTC<br>TGGAG<br>AGTAAGT<br>TTATCCAC<br>ACTACCA<br>ATGCCTA<br>GATTCTAT<br>TTAGC<br>GCTTCTGC<br>AGATCTG<br>GCATAGT<br>TATTAATT<br>GCAAATG<br>GGTAG<br>GTTGICTG<br>AATAGCAC<br>AGAAAAA<br>TGGAGATC<br>AGGTATTC<br>AAGCAT |
| 412.             | H3015B08-3 | BG064069       | ESTs<br>BG064069                                                                                                                                                                   | H3015B08      | Mm.197224     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 413.             | C0842H05-3 | Fbln1          | fibulin 1                                                                                                                                                                          | C0842H05      | Mm.219663     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 414.             | G0117D07-3 | Otx2           | orthodenticle<br>homolog 2<br>(Drosophila)                                                                                                                                         | G0117D07      | Mm.134516     | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 415.             | L0806E03-3 | Stmn4          | stathmin-like 4                                                                                                                                                                    | L0806E03      | Mm.35474      | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 416.             | H3073B06-3 | BG069137       | ESTs<br>BG069137                                                                                                                                                                   | H3073B06      | Mm.173781     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 417.             | H3082G08-3 | Myo10          | myosin X                                                                                                                                                                           | H3082G08      | Mm.60590      | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 418.             | C0141F07-3 | C3ar1          | complement<br>component 3a<br>receptor 1                                                                                                                                           | C0141F07      | Mm.2408       | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 419.             | K0525G09-3 | 5830411I20     | hypothetical<br>protein<br>5830411I20                                                                                                                                              | K0525G09      | Mm.31672      | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 420.             | H3064D01-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>ref:NP_001362.<br>1 (H.sapiens)<br>dynein,<br>axonemal,<br>heavy<br>polypeptide 8<br>[Homo sapiens]<br>RIKEN cDNA | H3064D01      | Mm.250054     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 421.             | C0120F08-3 | 6330406L22Rik  |                                                                                                                                                                                    | C0120F08      | Mm.5202       | Chromosome 10            | TAAAGGAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                      | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------|------------|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 6330406L22<br>gene                                                                                                                                                                                |               |               |                          | TCCACAT<br>TCACAAT<br>TTGAAAT<br>TTCTTAA<br>CTGCC<br>GTCACTT<br>GGTGTAT<br>GCACAAA<br>GCTCAGA<br>AAAGTTC<br>GTGAAC<br>GTCATGA<br>AATACAC<br>GAAATGT<br>TTCTTTAT<br>AAACGTC<br>GTTCA<br>TGTCGAT<br>TCTAAAG<br>CAACTTC<br>CATAGGG<br>TCATATA<br>CATTT<br>ATGCAAA<br>AAAAGCA<br>AAAAAAT<br>ATTGGAC<br>GAAGAGT<br>CCAAGCA<br>TTTGAGAC<br>TTCATAAC<br>AATTATAC<br>TATCCAA<br>ACTGCAA<br>TGGAG<br>ACCTAAA<br>CACAGGC<br>TACTTTGT<br>TAAATTGT<br>ATCATATC<br>TGCCC |
| 422.             | H3105G04-3 | Map4k4         | mitogen-<br>activated<br>protein kinase<br>kinase kinase 4                                                                                                                                        | H3105G04      | Mm.987        | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 423.             | J0800D09-3 | 2310004L02Rik  | RIKEN cDNA<br>2310004L02<br>gene                                                                                                                                                                  | J0800D09      | Mm.159956     | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 424.             | L0226H02-3 | 5830411I20     | hypothetical<br>protein<br>5830411I20                                                                                                                                                             | L0226H02      | Mm.31672      | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 425.             | L0529D10-3 | BM123730       | ESTs<br>BM123730                                                                                                                                                                                  | L0529D10      | Mm.221754     | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 426.             | H3088E05-3 | Gla            | galactosidase,<br>alpha                                                                                                                                                                           | H3088E05      | Mm.1114       | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 427.             | K0621H11-3 |                | K0621H11-3<br>NIA Mouse<br>Hematopoietic<br>Stem Cell (Lin-<br>/c-Kit-/Sca-1+)<br>cDNA Library<br>(Long) Mus<br>musculus<br>cDNA clone<br>NIA:K0621H11<br>IMAGE:30070<br>846 3', MRNA<br>sequence | K0621H11      |               | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 428.             | C0846H03-3 | D330025I23Rik  | RIKEN cDNA<br>D330025I23<br>gene                                                                                                                                                                  | C0846H03      | Mm.260376     | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 429.             | J0058E06-3 | C78984         | ESTs C78984                                                                                                                                                                                       | J0058E06      | Mm.249886     | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 430.             | K0325E09-3 | Ibsp           | integrin binding<br>sialoprotein                                                                                                                                                                  | K0325E09      | Mm.4987       | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                    | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------|------------|---------------|------------------------------------------------------------------------------|------------|------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 431.       | K0336F07-3 | Pycs          | pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase) | K0336F07   | Mm.233117  | Chromosome 19                         | TTTGACT<br>CCAGCCC<br>ATTCTCA<br>CTCGACA<br>ATTTCAT<br>TTTAC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 432.       | H3013B04-3 | B230106I24Rik | RIKEN cDNA B230106I24 gene                                                   | H3013B04   | Mm.24576   | Chromosome 3                          | AGGACTC<br>ACTTACA<br>GATGCCG<br>GAATGTT<br>GCATGAC<br>TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG |
| 433.       | L0238A07-3 | Midn          | midnolin                                                                     | L0238A07   | Mm.143813  | No Chromosome location info available | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 434.       | L0929C04-3 | Tnfrsf11b     | tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)     | L0929C04   | Mm.15383   | Chromosome 15: not placed             | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 435.       | L0020F05-3 | 6330583M11Rik | RIKEN cDNA 6330583M11 gene                                                   | L0020F05   | Mm.23572   | Chromosome 2                          | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 436.       | H3012H07-3 | Cd44          | CD44 antigen                                                                 | H3012H07   | Mm.24138   | Chromosome 2                          | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 437.       | K0240E11-3 | Myo5a         | myosin Va                                                                    | K0240E11   | Mm.3645    | Chromosome 9                          | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 438.       | K0401C06-3 | Col8a1        | procollagen, type VIII, alpha 1                                              | K0401C06   | Mm.86813   | Chromosome 16                         | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 439.       | C0917F02-3 | Frzb          | frizzled-related protein                                                     | C0917F02   | Mm.136022  | Chromosome 2                          | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 440.       | H3104C03-3 | 1500015O10Rik | RIKEN cDNA 1500015O10 gene                                                   | H3104C03   | Mm.11819   | Chromosome 1                          | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |
| 441.       | K0438D09-3 | Col8a1        | procollagen, type VIII, alpha 1                                              | K0438D09   | Mm.86813   | Chromosome 16                         | TAACC<br>CCACCTC<br>AAGTCTC<br>TACTGAA<br>AATTTGA<br>AAGAGAA<br>ATTTAC<br>GATGTTC<br>GTAAAAG<br>TAATATA<br>TAAGACT<br>CAGTATT<br>TTTAT<br>CTTAAGA<br>GGAAAAT<br>CTTTCTGC<br>TCCTAGC<br>ACAGAAC<br>CTCCGA<br>TATATTG<br>CCATAAC<br>AAAACCTG<br>TTTAGCT<br>TCGACCC<br>CTGTC<br>TCITTAGT<br>GCATTTA<br>GCATACA<br>ACAATCC<br>TGTATGA<br>TTGTG<br>AATCTATC<br>GATACTG<br>TTCTACCA<br>TGCTAAT<br>GAGCTAA<br>ATACTC<br>AATTTACA<br>GTGGTAG<br>AGGTCCAC<br>CCTAAGTT<br>GTGTGCTC<br>AATAA<br>ATGAGGC<br>ATTTGAAC<br>ATGTCAAC<br>TGGCTAAA<br>AAATCGA<br>GGCCATG<br>TCTACTAC<br>GCTTATCA<br>TCACTGCA<br>GAGGCAAC<br>ATGGGTTG                                                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                       |
|------------------|------------|----------------|---------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 442.             | H3152C04-3 | Usp16          | ubiquitin<br>specific<br>protease 16                                                        | H3152C04      | Mm.196253     | Chromosome 16                            | TCTTCA<br>GTACTGA/<br>ACAAGC/<br>TCCTATT<br>GAGAGA/<br>TGTGATA<br>AAAGTG<br>TTGGCCC<br>CCAAAGC/<br>AAGATT/<br>TAAATA/<br>CTGTATA<br>GTGCTT<br>CTGGGA/<br>CTAATGC<br>ATTCTG<br>CATTTAT<br>ACCTTAT<br>CTATT<br>TCCTCTG<br>AAATGAC<br>ACCTTGT<br>ATGGAG/<br>CAAAAGC/<br>GATTTT<br>GCCGCA/<br>AACAGA/<br>GTTTTTA<br>CATGTAA<br>AGGGATC<br>TCAACCC                                         |
| 443.             | H3079D12-3 | Pld3           | phospholipase<br>D3                                                                         | H3079D12      | Mm.6483       | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                       |
| 444.             | L0020E08-3 | Clqg           | complement<br>component 1, q<br>subcomponent,<br>gamma<br>polypeptide                       | L0020E08      | Mm.3453       | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                       |
| 445.             | J0025G01-3 | Yars           | tyrosyl-tRNA<br>synthetase                                                                  | J0025G01      | Mm.22929      | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                       |
| 446.             | L0832H09-3 | Mafb           | v-maf<br>musculoaponeu<br>rotic<br>fibrosarcoma<br>oncogene<br>family, protein<br>B (avian) | L0832H09      | Mm.233891     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                       |
| 447.             | C0451C02-3 | 2700094L05Rik  | RIKEN cDNA<br>2700094L05<br>gene                                                            | C0451C02      | Mm.25941      | No Chromosome location<br>info available | ACTTTTG/<br>TTTAGAA<br>GCCCACC<br>GAGTCTC<br>TCTGTTG/<br>GACCT<br>TGCTTAC/<br>AAGCCAC<br>GGTGGGI<br>GCTCTCT/<br>GAAGGA/<br>GCTTCT<br>TCCCAAT/<br>AGAAATC<br>ATGTAAC<br>TGGTACA<br>CACTGGA<br>ATAGA<br>CTTATGG<br>TATGTCC<br>GAATTCA<br>AAAACCTG<br>AAACCCCT<br>GAGTCA<br>GCCATAT<br>AACAGAA<br>AAGAATG<br>TTTATGC/<br>TAACCTC/<br>GCAGT<br>TCATTTTC<br>TCTAGGC/<br>GCTAAAC<br>ACTATGG |
| 448.             | H3063A08-3 | Lgmn           | legumain                                                                                    | H3063A08      | Mm.17185      | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                       |
| 449.             | K0629D05-3 | Evi2a          | ecotropic viral<br>integration site<br>2a                                                   | K0629D05      | Mm.3266       | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                       |
| 450.             | G0111D11-3 | Ctsl           | cathepsin L                                                                                 | G0111D11      | Mm.930        | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                       |
| 451.             | H3077D05-3 | Npc2           | Niemann Pick<br>type C2                                                                     | H3077D05      | Mm.29454      | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                       |
| 452.             | G0104C04-3 | Dab2           | disabled<br>homolog 2<br>(Drosophila)                                                       | G0104C04      | Mm.288252     | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                       |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                     | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60m<br>SEQUENCE                                                                  |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------|---------------|---------------|---------------------------------|----------------------------------------------------------------------------------|
| 453.             | L0502D10-3 | Pla1a          | phospholipase<br>A1 member A                                                                     | L0502D10      | Mm.24223      | Chromosome 16                   | ACGTAA/<br>GCTCC<br>CAACATC<br>GCTTTAC<br>ATGCCCT<br>GCTTCTC<br>TCGACAC<br>GTGAT |
| 454.             | H3126B08-3 | Pla2g7         | phospholipase<br>A2, group VII<br>(platelet-<br>activating factor<br>acetylhydrolase,<br>plasma) | H3126B08      | Mm.9277       | Chromosome 17                   | TTACCCA<br>AGCATT<br>AATATAC<br>TACTGTA<br>AGTGATC<br>GCCTAG                     |
| 455.             | J0034A07-3 | Creg           | cellular<br>repressor of<br>E1A-stimulated<br>genes                                              | J0034A07      | Mm.459        | Chromosome 1                    | ATAAGCC<br>CTGGGTC<br>ACTACTT<br>GGACCTA<br>AGTGACA<br>AAGAA                     |
| 456.             | H3114B07-3 | Slc12a4        | solute carrier<br>family 12,<br>member 4                                                         | H3114B07      | Mm.4190       | Chromosome 8                    | AAGTGG/<br>GAGCCGC<br>AGCTGAC<br>ACTTTTT<br>TAAAAACA<br>GTACTTC                  |
| 457.             | K0339H12-3 | Thbs1          | thrombospondi<br>n 1                                                                             | K0339H12      | Mm.4159       | Chromosome 2                    | CTTAAAA<br>TGTTGTG<br>AAAAGTC<br>GTTGTAC<br>CATAAAA<br>TTTGCC                    |
| 458.             | H3028C09-3 | Adk            | adenosine<br>kinase                                                                              | H3028C09      | Mm.19352      | Chromosome 14                   | CAGCTGC<br>CCCGCAA<br>TGCATTAC<br>CAGACTG<br>CTGCTTA<br>TG GTA                   |
| 459.             | L0277B06-3 | Psap           | prosaposin                                                                                       | L0277B06      | Mm.233010     | Chromosome 10                   | CTGTGGT<br>AGGAGTT<br>TGGATGA<br>AAGCACA<br>TGATCAG<br>TTAGAG                    |
| 460.             | H3013F05-3 | Sdc1           | syndecan 1                                                                                       | H3013F05      | Mm.2580       | Chromosome Multiple<br>Mappings | TTGTTTT<br>TTAACCT<br>GAACCAA<br>GGACGCC<br>ACGTAGG<br>GTTTG                     |
| 461.             | H3084A06-3 | Spin           | spindlin                                                                                         | H3084A06      | Mm.42193      | Chromosome 13                   | TGCCTGA<br>ACTTAAC<br>ATTGTCT<br>GATGAAA<br>TCCAAAG<br>CACAG                     |
| 462.             | H3077F04-3 | Osbpl8         | oxysterol<br>binding protein-<br>like 8                                                          | H3077F04      | Mm.134712     | Chromosome 10                   | ACTTCAG<br>TGGGTTT<br>AGTCAAG<br>GGCATTG<br>GTTTTGT<br>TAGGA                     |
| 463.             | K0324A06-3 | Itga11         | integrin, alpha<br>11                                                                            | K0324A06      | Mm.34883      | Chromosome 9                    | TCCCCTA<br>GTACGAC<br>ACTGTCA<br>TATATTTA                                        |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                            | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]        | 60m SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------|------------|---------------|----------------------------------------------------------------------|------------|------------|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 464.       | C0115E05-3 | 2010110K16Rik | RIKEN cDNA 2010110K16 gene                                           | C0115E05   | Mm.9953    | Chromosome 9                 | AAATGTT<br>ACGGT<br>GATCCAC<br>CTATGAA<br>GCAAAC<br>GTATCTC<br>AAAGGG<br>ATTCAGA<br>CATGACT<br>AGTTCTC<br>TCACAAA<br>TTACATG<br>TTCATGT<br>CTTGG<br>CTTGTA<br>ACACGTC<br>CCTAAAA<br>GGGTATA<br>AAAATTT<br>CCATGG<br>TGTCTGA<br>GCTTGAA<br>TCAACCA<br>CCAGTTC<br>CAGACTT<br>CATAT<br>TACTCCC<br>CTATTTG<br>TAATAGT<br>CGCCACA<br>TACTGTT<br>TTCAA<br>CAGCCGA<br>TTTTCAA<br>ATTTTAA<br>TTGTGTA<br>AACCAAG<br>AAGAG<br>GGGACAC<br>TTTACAT<br>TTTAACC<br>GAAAGAC<br>AGATAGA<br>AAGACAC<br>GCCTGCC<br>ACCCAG<br>GAGTCTA<br>CAAAAAC<br>CAAACCTC<br>TTTTTAA<br>AATCTAG<br>TAGAAAT<br>GTGTATG<br>ATTGTAT<br>ACCATA<br>GACCG<br>ACGATGA<br>GTGTTTG<br>CTTTCCAC<br>GAACTAT<br>CGGAAAA<br>AATGTTT<br>GATGCGT<br>ATGTTCCI<br>GGAAAAAG<br>TTCAAGCC |
| 465.       | C0668G11-3 | Fabp5         | fatty acid binding protein 5, epidermal                              | C0668G11   | Mm.741     | Chromosome Multiple Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 466.       | L0030A03-3 | Alox5ap       | arachidonate 5-lipoxygenase activating protein                       | L0030A03   | Mm.19844   | Chromosome 5                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 467.       | H3009E11-3 | Socs3         | suppressor of cytokine signaling 3                                   | H3009E11   | Mm.3468    | Chromosome 11                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 468.       | L0010B01-3 | Abca1         | ATP-binding cassette, sub-family A (ABC1), member 1                  | L0010B01   | Mm.369     | Chromosome 4                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 469.       | G0116C07-3 | Ctsb          | cathepsin B                                                          | G0116C07   | Mm.22753   | Chromosome 14                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 470.       | K0426E09-3 | Eps8          | epidermal growth factor receptor pathway substrate 8                 | K0426E09   | Mm.2012    | Chromosome 6                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 471.       | H3102F08-3 | Asah1         | N-acylsphingosine amidohydrolase 1                                   | H3102F08   | Mm.22547   | Chromosome 8                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 472.       | L0825G08-3 | Dcamk1l       | double cortin and calcium/calmodulin-dependent protein kinase-like 1 | L0825G08   | Mm.39298   | Chromosome 3                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 473.       | K0306B10-3 | Fgf7          | fibroblast growth factor 7                                           | K0306B10   | Mm.57177   | Chromosome 2                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 474.       | H3127F04-3 | Chst11        | carbohydrate sulfotransferase 11                                     | H3127F04   | Mm.41333   | Chromosome 10                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                        | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                   |
|------------------|------------|----------------|---------------------------------------------------------------------|---------------|---------------|------------------------------------------|-----------------------------------------------------------------------------------|
| 475.             | L0208A08-3 | 1200013B22Rik  | RIKEN cDNA<br>1200013B22<br>gene                                    | L0208A08      | Mm.100666     | Chromosome 1                             | TTATTTT<br>AGTAAC<br>CATCTTA<br>TCAGAGA<br>AACCTTG<br>TGTTCTT<br>CCCAGAT<br>ATGGA |
| 476.             | H3026G09-3 | Col2a1         | procollagen,<br>type II, alpha 1                                    | H3026G09      | Mm.2423       | Chromosome 15                            | CGTGTCC<br>CAATGGT<br>TTCTGTG<br>ACACCTC<br>TTTTTTA<br>ATCAA                      |
| 477.             | C0218D02-3 | Madh1          | MAD homolog<br>1 (Drosophila)                                       | C0218D02      | Mm.15185      | Chromosome 8                             | AAGGAGC<br>GATAATA<br>ACCTCTG<br>CAACTAT<br>TTGAGAA<br>ACAAGC                     |
| 478.             | J1031F04-3 | Dfna5h         | deafness,<br>autosomal<br>dominant 5<br>homolog<br>(human)          | J1031F04      | Mm.20458      | No Chromosome location<br>info available | GTTTATA<br>GACCTAA<br>ATAAAAC<br>GGGTATC<br>TAACGTT<br>AAAAGA                     |
| 479.             | L0276A08-3 | Rai14          | retinoic acid<br>induced 14                                         | L0276A08      | Mm.26786      | Chromosome 15                            | AAACTTG<br>ATTTTGT<br>CGCCTGA<br>GCGTAGC<br>TTCTTGT<br>GGATG                      |
| 480.             | C0508H08-3 | Sptlc2         | serine<br>palmitoyltransf<br>erase, long<br>chain base<br>subunit 2 | C0508H08      | Mm.565        | Chromosome 12                            | CTCATAC<br>GAAATAC<br>CACTGCT<br>AGGAGAT<br>TGAAGTT<br>ATCTGC                     |
| 481.             | J0042D09-3 | C78076         | ESTs C78076                                                         | J0042D09      | Mm.290404     | Chromosome 12                            | AAATCCA<br>TTAAAAG<br>GTTTCTT<br>TAAGTGA<br>CATTACT<br>TATAC                      |
| 482.             | J0013B06-3 | Akr1b8         | aldo-keto<br>reductase<br>family 1,<br>member B8                    | J0013B06      | Mm.5378       | Chromosome 6                             | ACCAGGA<br>TGGTAAC<br>GAGGGCA<br>AGATAAA<br>ATAAAGA<br>AGAACAT                    |
| 483.             | H3158D11-3 | Mmp2           | matrix<br>metalloproteinase 2                                       | H3158D11      | Mm.29564      | Chromosome 8                             | TCAACAT<br>GACCTTT<br>GGTTTCA<br>TCTCAGA<br>ATAGAGA<br>GCTTAG                     |
| 484.             | H3001D04-3 | Hist2h3c2      | histone 2, H3c2                                                     | H3001D04      | Mm.261624     | Chromosome 13                            | GACCGAG<br>CACCACA<br>CCAAGGG<br>ATAAGAC<br>CCGTTCA<br>CCCGAAA                    |
| 485.             | C0664G04-3 | Ppicap         | peptidylprolyl<br>isomerase C-<br>associated<br>protein             | C0664G04      | Mm.3152       | Chromosome 11                            | TTCTACCT<br>TAACTCC<br>ACATGGT<br>ATGGTAC                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                  | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------|------------|----------------|-----------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 486.             | H3091E10-3 | Nupr1          | nuclear protein<br>1                                                                          | H3091E10      | Mm.18742      | Chromosome 7                             | CAGTGGT<br>ATGCA<br>TTGGAGA<br>AGGAGT<br>GCAGGAC<br>GGCCTGC<br>TTCTTTC<br>CTAAGT<br>TTATTGA<br>TTTGAAG<br>AACTTAG<br>TTGGAAT<br>GCATAAA<br>GACTGC<br>CACCATT<br>ACTTGCT<br>CACTAAT<br>CTGCATT<br>GCAACAA<br>ATGTTT<br>AACAAAG<br>CCTGTGG<br>GGGGGTC<br>TAAGTTA<br>CCAATAA<br>TTACCT<br>TTTTGAC<br>TGAACCC<br>TGTTTTCC<br>CGAACAC<br>ATAATAT<br>AAAGC<br>GTGAGGA<br>AATTAGA<br>TCATAAG<br>ATATGAC<br>CATTCTT<br>ATGACC<br>CGCCCTG<br>CTCTGTC<br>CTTGGA<br>TAAAAAT<br>GCTTTTTC<br>CAGCAA<br>AAGATGG<br>GTTGTCC<br>AAGATCC<br>GTCTAAA<br>GCAAGGG<br>TGAGGTG<br>GTTTTAA<br>TGCCAGG<br>CATTTTTC<br>TGAAACC<br>GATGTTT<br>AACAC<br>TCTGAGG<br>AAAATAT<br>ACTGAAT<br>CCAAATG<br>AGGGAGA<br>TTCCTG<br>AAGTATT<br>GACTGAA<br>ACTTGAA<br>TCAGAGA |
| 487.             | X98792.1   | Ptgs2          | prostaglandin-<br>endoperoxide<br>synthase 2                                                  | X98792        | Mm.3137       | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 488.             | L0908B12-3 | Ptpn1          | protein tyrosine<br>phosphatase,<br>non-receptor<br>type 1                                    | L0908B12      | Mm.227260     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 489.             | H3081D02-3 | Bok            | Bcl-2-related<br>ovarian killer<br>protein                                                    | H3081D02      | Mm.3295       | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 490.             | C0127E12-3 | Cln5           | ceroid-<br>lipofuscinosis,<br>neuronal 5                                                      | C0127E12      | Mm.38783      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 491.             | K0310G10-3 | Col5a2         | procollagen,<br>type V, alpha 2                                                               | K0310G10      | Mm.257899     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 492.             | H3023H09-3 | Ftl1           | ferritin light<br>chain 1                                                                     | H3023H09      | Mm.7500       | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 493.             | G0104B11-3 | Slc7a7         | solute carrier<br>family 7<br>(cationic amino<br>acid transporter,<br>y+ system),<br>member 7 | G0104B11      | Mm.142455     | Chromosome 14                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 494.             | C0123F05-3 | B4galt5        | UDP-<br>Gal:betaGlcNA<br>c beta 1,4-<br>galactosyltransf<br>erase,<br>polypeptide 5           | C0123F05      | Mm.200886     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 495.             | H3082D01-3 | 1810015C04Rik  | RIKEN cDNA<br>1810015C04<br>gene                                                              | H3082D01      | Mm.25311      | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 496.             | C0121E07-3 | AW539579       | EST<br>AW539579                                                                               | C0121E07      | Mm.282049     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                     | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------|------------|----------------|--------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 497.             | H3153H08-3 | Hs6st2         | heparan sulfate<br>6-O-<br>sulfotransferase<br>2 | H3153H08      | Mm.41264      | Chromosome X             | AGACTGA/<br>AAGGTGT/<br>ACATTTT/<br>ATCATCA/<br>AATCCCA/<br>TTCAAAC/<br>AACATCT/<br>AGTGG/<br>CTGGGGA/<br>GATCTTT/<br>TTTGAAA/<br>ATAAGGA/<br>TCTGGTT/<br>TCTCAC/<br>AGGACTC/<br>ACTATAT/<br>CTGCTCT/<br>TAATGTT/<br>AAGCTCC/<br>GAAAGCC/<br>GCTCCAA/<br>CCATGTA/<br>ATAGACT/<br>CTACAAT/<br>ATAACGT/<br>AGCTT/<br>CAGCTGA/<br>GTTTTGG/<br>CAGGAAA/<br>GTCCAGA/<br>TGAAAAG/<br>CTAAGA/<br>TGTTTTTA/<br>TGTTTGG/<br>GAAGAAT/<br>ACACTTC/<br>CTAAATC/<br>AGCCCC/<br>TCCAGTT/<br>AAGAAGC/<br>TAGGAAT/<br>CTTGTGC/<br>ACTACAC/<br>ATGCTA/<br>CATAAAG/<br>AGTGGAG/<br>TGTTTAC/<br>CCGAATG/<br>GCTGAAC/<br>AGAAT/<br>GGATTTC/<br>GATGAAT/<br>GCACTTT/<br>ACTGCGG/<br>CAGTTAC/<br>ACACCC/<br>TGCTTTTA/<br>TGTTCTCG/<br>TTCCTGAA/<br>AGAGCCT/<br>GATAGTTC/<br>TGCAA/<br>TGAAGCA/<br>AACATAA/<br>CTCACCAC/<br>CTGCTGAA/ |
| 498.             | J0238C08-3 | 4930579A11Rik  | RIKEN cDNA<br>4930579A11<br>gene                 | J0238C08      | Mm.24584      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 499.             | L0942B10-3 | Ms2            | macrophage<br>scavenger<br>receptor 2            | L0942B10      | Mm.45173      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 500.             | J0915B05-3 | Cdca1          | cell division<br>cycle associated<br>1           | J0915B05      | Mm.151315     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 501.             | H3058B09-3 | Lypla3         | lysophospholip<br>ase 3                          | H3058B09      | Mm.25492      | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 502.             | C0197E01-3 | D630023B12     | hypothetical<br>protein<br>D630023B12            | C0197E01      | Mm.227732     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 503.             | J0802G04-3 | 0610011I04Rik  | RIKEN cDNA<br>0610011I04<br>gene                 | J0802G04      | Mm.27061      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 504.             | H3039E08-3 | Sh3d3          | SH3 domain<br>protein 3                          | H3039E08      | Mm.4165       | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 505.             | L0210A08-3 | B130023O14Rik  | RIKEN cDNA<br>B130023O14<br>gene                 | L0210A08      | Mm.27098      | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 506.             | H3114C10-3 | Ppgb           | protective<br>protein for beta-<br>galactosidase | H3114C10      | Mm.7046       | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 507.             | C0322A01-3 | 2810441C07Rik  | RIKEN cDNA<br>2810441C07<br>gene                 | C0322A01      | Mm.29329      | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                            | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]    | 60m SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------|------------|---------------|----------------------------------------------------------------------|------------|------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 508.       | L0256F11-3 | Adfp          | adipose differentiation related protein                              | L0256F11   | Mm.381     | Chromosome 4             | AGAACCT<br>TTGGGGC<br>GAATCCT<br>TGAAGT<br>ATTACTT<br>AACAAAC/<br>CTCTCAA<br>CTGGTA<br>GATATTA<br>TATATCA<br>ACTTGAC<br>AAAGATC<br>CACCCCC<br>TGTTG<br>TGTGATA<br>TGTGACA<br>TATTAGT<br>ACATATT<br>CTCCAAA<br>TTTGC<br>TAAAGT<br>GTAAGCC<br>AAAGGA/<br>GTATCTA<br>GCTTTCC.<br>TAATCAG<br>GGAGATT<br>TCTTCAG<br>TCTACAT.<br>TACACAC<br>GTGTCTT.<br>AGCAA<br>AATCCAT<br>GGGGGG/<br>AGTCCAG<br>CTTAAGA<br>AGTAAAA<br>TGGCTT<br>AATGTGG<br>TGGAGAA<br>CATTTCT<br>TGATAAC<br>CCTGTTG<br>GACAGT<br>TGACATG<br>AAATCAA<br>ATTTTAC<br>AGAAGTA<br>AATCTCT<br>GCCAAGC<br>ACTGGAT<br>TAACAT<br>ATAAAAT<br>AAGTGAC<br>CGTCTAC.<br>TTCCAG<br>ATACAAG<br>GCTGTTA.<br>TCTTGGA<br>ATTCTAT<br>TGTATAC<br>ATCAAC<br>AGCATCA<br>TCCTGTC/<br>ACAAAAA<br>AAGAAGA |
| 509.       | L0939H06-3 | Mgat5         | mannoside acetylglucosaminyltransferase 5                            | L0939H06   | Mm.38399   | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 510.       | C0503B05-3 | Dcamk1l       | double cortin and calcium/calmodulin-dependent protein kinase-like 1 | C0503B05   | Mm.39298   | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 511.       | H3136H11-3 | Map4k5        | mitogen-activated protein kinase kinase 5                            | H3136H11   | Mm.260244  | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 512.       | K0349A04-3 | Fn1           | fibronectin 1                                                        | K0349A04   | Mm.193099  | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 513.       | C0177C04-3 | Ctsz          | cathepsin Z                                                          | C0177C04   | Mm.156919  | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 514.       | C0668D08-3 | Grn           | granulin                                                             | C0668D08   | Mm.1568    | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 515.       | C0106D12-3 | Anxa1         | annexin A1                                                           | C0106D12   | Mm.14860   | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 516.       | H3078E09-3 | Hexb          | hexosaminidase B                                                     | H3078E09   | Mm.27816   | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 517.       | L0033F05-3 | 2810442I22Rik | RIKEN cDNA 2810442I22 gene                                           | L0033F05   | Mm.275696  | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 518.       | K0144G04-3 | Ifi203        | interferon activated gene 203                                        | K0144G04   | Mm.245007  | Chromosome 1: not placed |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                 | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A] | 60m SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------|------------|---------------|---------------------------------------------------------------------------|------------|------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 519.       | H3144E05-3 | 4933426M11Rik | RIKEN cDNA 4933426M11 gene                                                | H3144E05   | Mm.27112   | Chromosome 12         | TAATTAC<br>AAGATGC<br>CCTCTGT<br>AGGAAC<br>AGCATAC<br>ATGGAA<br>TGCAAAC<br>CTAGAT<br>GTGTAGA<br>TATTGAA<br>CAGTCCT<br>AGACCA<br>TAATTCT<br>AATGG<br>CTGATCC<br>TCATCTC<br>CTCCGTG<br>CCTAGCA<br>AAGTCAA<br>GGTTT<br>TGTAGAA<br>TGGCCTC<br>TATAAAT<br>ATAAATG<br>ATTTAAT<br>GTTTC<br>GGTGCCA<br>AGAAGAC<br>AGTTGGA<br>ATACCCG<br>AATTCCA<br>TAGTCAA<br>CAGTGTT<br>AAGAGAA<br>AAAGTTC<br>GGTTTGG<br>GGATCAA<br>GGAAACA<br>ATAACTA<br>TACTTAG<br>TGTCTAA<br>TTGCCAC<br>ATTGGTC<br>CAGCA<br>CCTTGGG<br>TTTTGTG<br>AGTTTGC<br>AGATAAC<br>GCAATAA<br>TACAGCA<br>TCTATAC<br>ATAAAAA<br>ACCTACA<br>ACTGTAA<br>TCATGTT<br>GGCAAG<br>CCTGTTT<br>AACCCTC<br>TCTACCG<br>CGTGAAT<br>AAAGCCT<br>GAGTC<br>ACCGTGT<br>ACTCATA<br>GCATGAC<br>TCTACCA |
| 520.       | K0336D02-3 | Ifi16         | interferon, gamma-inducible protein 16                                    | K0336D02   | Mm.212870  | Chromosome 1          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 521.       | H3004B12-3 | Hpn           | hepsin                                                                    | H3004B12   | Mm.19182   | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 522.       | K0617G07-3 | Atp6v1b2      | ATPase, H <sup>+</sup> -transporting, V1 subunit B, isoform 2             | K0617G07   | Mm.10727   | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 523.       | L0849B10-3 | Pltp          | phospholipid transfer protein                                             | L0849B10   | Mm.6105    | Chromosome 2          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 524.       | L0019H03-3 | Fn1           | fibronectin 1                                                             | L0019H03   | Mm.193099  | Chromosome 1          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 525.       | J0099E12-3 | Slc6a6        | solute carrier family 6 (neurotransmitter transporter, taurine), member 6 | J0099E12   | Mm.200518  | Chromosome 6          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 526.       | J0023G04-3 | BC004044      | cDNA sequence BC004044                                                    | J0023G04   | Mm.6419    | Chromosome 5          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 527.       | C0913D04-3 | 4933433D23Rik | RIKEN cDNA 4933433D23 gene                                                | C0913D04   | Mm.46067   | Chromosome 14         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 528.       | H3020C02-3 | Mt1           | metallothionein 1                                                         | H3020C02   | Mm.192991  | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 529.       | C0217B11-3 | Sema4d        | sema domain, immunoglobulin domain (Ig), transmembrane                    | C0217B11   | Mm.33903   | Chromosome 13         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                              |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | domain (TM)<br>and short<br>cytoplasmic<br>domain,<br>(semaphorin)<br>4D                                                                             |               |               |                          | GTGTAAA<br>TGTGT                                                                                                                                                                                                                                                                                                                             |
| 530.             | C0917E01-3 | Bhlhb2         | basic helix-<br>loop-helix<br>domain<br>containing,<br>class B2                                                                                      | C0917E01      | Mm.2436       | Chromosome 6             | GCCAAAC<br>AATGTTT<br>TGTCTAT<br>ATAATTA<br>ATCTACC<br>GAGGAA<br>TCCAGAA<br>CATTGCC<br>TCACACC<br>AATTGTC<br>CATCGCT<br>GCATT<br>AAGGAC<br>GGCCATC<br>GTCAGTA<br>CATTACT<br>CCTCTCT<br>TGAAT<br>ATCTCCC<br>CAAAGA/<br>AAACTCA<br>CTGTCTG<br>GAAGAA/<br>GTGTTGT                                                                               |
| 531.             | H3132B12-5 | Deaf1          | deformed<br>epidermal<br>autoregulatory<br>factor 1<br>(Drosophila)                                                                                  | H3132B12      | Mm.28392      | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                              |
| 532.             | L0270C04-3 | Mpp1           | membrane<br>protein,<br>palmitoylated                                                                                                                | L0270C04      | Mm.2814       | Chromosome X             |                                                                                                                                                                                                                                                                                                                                              |
| 533.             | J0709H10-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>pir:A38712<br>(H.sapiens)<br>A38712<br>fibrillarin<br>[validated] -<br>human | J0709H10      | Mm.296913     | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                              |
| 534.             | C0166A10-3 | Car2           | carbonic<br>anhydrase 2                                                                                                                              | C0166A10      | Mm.1186       | Chromosome 3             | ATGAAGC<br>GATAATT<br>ACAAGTC<br>TCATGAG<br>ACTGAAC<br>TTAGGC<br>GGTGTAC<br>ACAATAC<br>AATACAA<br>ATATTCT<br>ACAATCT<br>GGTGTGC<br>GGAGAAC<br>ATTATCT<br>GGCTTCC<br>TCTGTTT<br>ACTGGTA<br>GTGGAC<br>GTGAACA<br>GAATTTA<br>CCATACT<br>CAGGTAC<br>ATTCTTC<br>CTCTAC<br>GGCTTCA<br>TGTGGAC<br>GCCCAA<br>AATGACC<br>TATATGT<br>GCCTCT<br>GTTTGTA |
| 535.             | L0511A03-3 | BM122519       | ESTs<br>BM122519                                                                                                                                     | L0511A03      | Mm.296074     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                              |
| 536.             | H3029F09-3 | Atp6v1e1       | ATPase, H+<br>transporting,<br>V1 subunit E<br>isoform 1                                                                                             | H3029F09      | Mm.29045      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                              |
| 537.             | J0716H11-3 | Kdt1           | kidney cell line<br>derived<br>transcript 1                                                                                                          | J0716H11      | Mm.1314       | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                              |
| 538.             | C0102C01-3 | Acp5           | acid<br>phosphatase 5,<br>tartrate resistant                                                                                                         | C0102C01      | Mm.46354      | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                              |
| 539.             | C0641C07-3 | Pdgfb          | platelet derived                                                                                                                                     | C0641C07      | Mm.144089     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                              |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                     | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------|------------|----------------|------------------------------------------------------------------|---------------|---------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | growth factor,<br>B polypeptide                                  |               |               |                          | TGGTGAT<br>TTTTTTG<br>CTTCTT<br>TTTTTAA<br>AAAG<br>ATGGAA<br>TTAGAGT<br>AAGAGA<br>CAGATAC<br>GGCTGGC<br>GAGGTC<br>AATAGTG<br>ATTGTG<br>CAGAATT<br>AGGTCAT<br>AATCCTT<br>GGGTAAC<br>TATGAAG<br>GGGAAAC<br>AGCTATC<br>ACCTGGA<br>CTCAGCC<br>TAACAGT<br>GAGGCAA<br>CCTTATT<br>AACTAGT<br>AAAGATT<br>TAAGCCC<br>GATGG<br>TAATGAA<br>GTATAAT<br>GCCAAAT<br>CTTGTTCT<br>GTCACGA<br>GTCTTG<br>CAGTTTGT<br>GTAGAAT<br>TTTCTAA<br>AAAAGCT<br>TTGAAGT<br>ACAGAG<br>TAGAAAA<br>CACCAAC<br>GGCCTCC<br>GTCATCC<br>ACTAAGA<br>GATTCTT<br>TATCTAAC<br>CAAGTCT<br>CATTAGC<br>AGAAAGTA<br>CCAAGTGT<br>CACCT<br>AAATTATC<br>TGGATAC<br>GGAACAT<br>AGGCACA<br>ATGAATA<br>AAATCC<br>AACTATTC<br>GTATATTT<br>AACACAG<br>ACTGTGG<br>TATCTGCT<br>AGCAA<br>ACCTCTGC |
| 540.             | C0147C09-3 | Ttc7           | tetratricopeptide repeat domain 7                                | C0147C09      | Mm.77396      | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 541.             | K0301G02-3 | 9430025M21Rik  | RIKEN cDNA 9430025M21 gene                                       | K0301G02      | Mm.87452      | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 542.             | H3022D05-3 | Tpbpb          | trophoblast specific protein beta                                | H3022D05      | Mm.297991     | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 543.             | H3007C09-3 | Sh3bgrl3       | SH3 domain binding glutamic acid-rich protein-like 3             | H3007C09      | Mm.22240      | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 544.             | L0820G02-3 | Igsf4          | immunoglobulin superfamily, member 4                             | L0820G02      | Mm.248549     | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 545.             | C0120H11-3 | 4933433D23Rik  | RIKEN cDNA 4933433D23 gene                                       | C0120H11      | Mm.46067      | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 546.             | J1016E08-3 | 1810046J19Rik  | RIKEN cDNA 1810046J19 gene                                       | J1016E08      | Mm.259614     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 547.             | L0822D10-3 | Prkcb          | protein kinase C, beta                                           | L0822D10      | Mm.4182       | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 548.             | H3050H09-3 | Ppp2r5c        | protein phosphatase 2, regulatory subunit B (B56), gamma isoform | H3050H09      | Mm.36389      | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 549.             | J0442H09-3 |                | Mus musculus hypothetical LOC237436 (LOC237436), mRNA            | J0442H09      | Mm.11982      | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 550.             | H3141E06-3 | Sra1           | steroid receptor                                                 | H3141E06      | Mm.29058      | Chromosome 18            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                 | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------|------------|----------------|------------------------------------------------------------------------------|---------------|---------------|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | RNA activator<br>1                                                           |               |               |                                 | AGGCAT<br>GGACTGC<br>GTCACAC<br>GAAACAC<br>TTTACA<br>CCAGTA1<br>ACAAAA<br>CCACAAC<br>CCGCATC<br>CAAGTTC<br>CCATAT<br>GTAAAGC<br>CATTACT<br>GTATTTC<br>GCATATT<br>TTAAGGC<br>TTCAAG<br>CTCTAAG<br>TCATTTT<br>AATTATT<br>AGAAATC<br>CTTATAC<br>GCAAT<br>TCTAATC<br>GGCCTTA<br>GTTCAGG<br>AGTAGAC<br>ATGCCAA<br>TCTTCTT<br>ATTCAGA<br>GAAAGG1<br>AATGGTC<br>TTACCAG<br>TCTACAT<br>TAATTT<br>CAGTTAT<br>TTCCATT<br>AATATCT<br>AACTGTA<br>CTATGAC<br>ACTGA<br>GCTTTCT<br>ACGTATT<br>AAATTGT<br>TGTGCCA<br>TCATGAT<br>GATGA<br>TGGCTAG<br>AATTGAG<br>AGGTTTC<br>AACCAGA<br>AAAAGCC<br>GTGTCG<br>AGAGGAC<br>ATGAAGA<br>TGTTCTC<br>CGGTCAG<br>AGCATAC<br>ACTGAAA<br>AGAAAA<br>AAAGCAG<br>AAAAGTT<br>GACATAG<br>CTGCTAA<br>GTCCTCT<br>ATATTTG |
| 551.             | C0170H06-3 | Adss2          | adenylosuccinat<br>e synthetase 2,<br>non muscle                             | C0170H06      | Mm.132946     | Chromosome 1                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 552.             | K0344C08-3 | Emp1           | epithelial<br>membrane<br>protein 1                                          | K0344C08      | Mm.30024      | Chromosome 6                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 553.             | J0907F03-3 | Npl            | N-<br>acetylneuramin<br>ate pyruvate<br>lyase                                | J0907F03      | Mm.24887      | Chromosome 1                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 554.             | J1008C10-3 | Ptpn1          | protein tyrosine<br>phosphatase,<br>non-receptor<br>type 1                   | J1008C10      | Mm.2668       | Chromosome 2                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 555.             | K0103F09-3 | 2500002K03Rik  | RIKEN cDNA<br>2500002K03<br>gene                                             | K0103F09      | Mm.29181      | Chromosome 6                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 556.             | C0837H01-3 | Adam9          | a disintegrin<br>and<br>metalloproteina<br>se domain 9<br>(meltrin<br>gamma) | C0837H01      | Mm.28908      | Chromosome 8                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 557.             | J0207H07-3 | Runx2          | runt related<br>transcription<br>factor 2                                    | J0207H07      | Mm.4509       | Chromosome 17                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 558.             | J0246C10-3 | Tpd52          | tumor protein<br>D52                                                         | J0246C10      | Mm.2777       | Chromosome Multiple<br>Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 559.             | H3158E12-3 | BC003324       | cDNA sequence<br>BC003324                                                    | H3158E12      | Mm.29656      | Chromosome 5                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 560.             | H3094A04-3 | Dnajc3         | DnaJ (Hsp40)<br>homolog,<br>subfamily C,<br>member 3                         | H3094A04      | Mm.12616      | Chromosome 14                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 561.             | L0231F01-3 | Evl            | Ena-vasodilator                                                              | L0231F01      | Mm.2144       | Chromosome 12                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                             | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------|------------|----------------|----------------------------------------------------------|---------------|---------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | stimulated<br>phosphoprotein                             |               |               |                          | TTTAAGC<br>GTTCCCTT<br>TTATAGA<br>ACCCCCA<br>ACCTG<br>GACTCTC<br>TTACAGA<br>TATCAGA<br>GAGAAG/<br>TGTTAAG<br>TTCACA<br>TAAAAATC<br>TGAAAGT<br>CTCAGTT/<br>GAATAAC<br>GTGTACC<br>TGGAATG<br>CCAATGA<br>ACAGTGT/<br>ACTTAAC<br>TCCAATA/<br>AATGCTT/<br>ATTTG<br>TCAAATC/<br>TCAACTT/<br>AAAATGG<br>TTTAATGC<br>GAGACTT/<br>GTCGG<br>CTATACAC<br>ATATGCT/<br>GATGTGA/<br>ATAATGG/<br>CTTTCCAC<br>GCACITT<br>CTGAGAT/<br>CAAATCTT<br>CAACTGAC<br>GGATGGA/<br>TTTAATTA<br>AACGG<br>AAATGTCT<br>CAACAGT/<br>GTACTATG<br>ATCCCTA<br>AAACTTCA<br>CAGCC<br>TGAACATT<br>AGGATTTC<br>CTATACTG<br>TAAACCCA<br>TTTTCTGG<br>CAGGG<br>CAACAAAC<br>ATTTACAT/<br>TAATCCAC<br>CTTAAAGA<br>ACAGTTAG<br>AGCAC<br>TGGACACA<br>CACTAAAT<br>TGATTTAG<br>AAGTAACT<br>ACTGAAAG<br>CCTAAAC<br>TTGTTGTG |
| 562.             | K0512E10-3 | Myo5a          | myosin Va                                                | K0512E10      | Mm.222258     | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 563.             | K0608H09-3 | Ptpnc          | protein tyrosine<br>phosphatase,<br>receptor type, C     | K0608H09      | Mm.143846     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 564.             | L0842E04-3 | Prkcb          | protein kinase<br>C, beta                                | L0842E04      | Mm.4182       | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 565.             | H3121G01-3 | BG073361       | ESTs<br>BG073361                                         | H3121G01      | Mm.182649     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 566.             | C0947F04-3 | 5830411K21Rik  | RIKEN cDNA<br>5830411K21<br>gene                         | C0947F04      | Mm.160141     | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 567.             | H3009D03-5 | Plac8          | placenta-<br>specific 8                                  | H3009D03      | Mm.34609      | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 568.             | H3132E07-3 | Lxn            | latexin                                                  | H3132E07      | Mm.2632       | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 569.             | H3054C01-3 | Nr2e3          | nuclear receptor<br>subfamily 2,<br>group E,<br>member 3 | H3054C01      | Mm.9652       | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 570.             | H3013H03-3 | Man1a          | mannosidase 1,<br>alpha                                  | H3013H03      | Mm.117294     | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 571.             | J0058F02-3 | ank            | progressive<br>ankylosis                                 | J0058F02      | Mm.142714     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 572.             | L0829D10-3 | Snca           | synuclein, alpha                                         | L0829D10      | Mm.17484      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                             | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                     |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 573.             | H3037H02-3 | 1110018O12Rik  | RIKEN cDNA<br>1110018O12<br>gene                                                                                                                                                         | H3037H02      | Mm.28252      | Chromosome 18                            | TCACACT<br>TTGTTAG<br>AACTTA/<br>CCTAAG/<br>ACCAC<br>TGAACA/<br>AGTATTC<br>GCTTCAC<br>AGTTAA/<br>AGTGAC/<br>ATGGAA<br>AAGGTC/<br>ATACAG/<br>TTTGCTA<br>CTAGAA/<br>CCATAA/<br>ACTGCA<br>GACTGA/<br>AAAGTT/<br>AACGGT/<br>CTCTAG/<br>TGTGGA/<br>TGATAT                                                                                                                                                                |
| 574.             | K0105H12-3 | Cdk6           | cyclin-<br>dependent<br>kinase 6                                                                                                                                                         | K0105H12      | Mm.88747      | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 575.             | C0105D10-3 |                | C0105D10-3<br>NIA Mouse<br>E7.5<br>Extraembryonic<br>Portion cDNA<br>Library Mus<br>musculus<br>cDNA clone<br>C0105D10 3',<br>MRNA<br>sequence<br>putative<br>serine/threonine<br>kinase | C0105D10      |               | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 576.             | L0229E05-3 | Prkx           |                                                                                                                                                                                          | L0229E05      | Mm.106185     | Chromosome X                             | TCAAAT/<br>AACCCCT/<br>AGGCTG/<br>CAAATG/<br>ATGCGA/<br>CTACAG<br>GCACTA/<br>TTCATCT<br>AAGGTT/<br>CTACAA/<br>CAAAAA/<br>ACAGGC<br>CTTGCA/<br>GCGTGT/<br>TTCTCGC<br>TCCTGAC<br>TGGAGT/<br>TGTTA<br>AGTGTT/<br>CAAAGC/<br>AAGCTC/<br>TGGTTA/<br>TGATTCT<br>CGTTCC<br>TCCAGA/<br>CAGAGA<br>GATCTTC<br>ATTTTC/<br>GTGCTA/<br>AAATTC<br>AGTGAC/<br>CCTTTTA/<br>CATTAA<br>GGAGCT<br>TAAAAG<br>ATTCCA<br>ACATAC<br>ATCACC |
| 577.             | L0931H07-3 |                | ESTs<br>BQ557106                                                                                                                                                                         | L0931H07      | Mm.221935     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 578.             | K0138B11-3 | Trim25         | tripartite motif<br>protein 25                                                                                                                                                           | K0138B11      | Mm.4973       | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 579.             | H3019H03-3 | Lass6          | longevity<br>assurance<br>homolog 6 (S.<br>cerevisiae)                                                                                                                                   | H3019H03      | Mm.265620     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 580.             | J0051F04-3 | Ifi30          | interferon<br>gamma<br>inducible<br>protein 30                                                                                                                                           | J0051F04      | Mm.30241      | Chromosome 8                             |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 581.             | H3106G04-3 | Cacna1d        | calcium<br>channel,<br>voltage-<br>dependent, L<br>type, alpha 1D<br>subunit                                                                                                             | H3106G04      | Mm.9772       | Chromosome 14                            |                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 582.             | L0701D10-3 | Arhgdib        | Rho, GDP<br>dissociation                                                                                                                                                                 | L0701D10      | Mm.2241       | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                     | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                         |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | inhibitor (GDI)<br>beta                                                                                                                                                          |               |               |                          | GTTTTAT<br>CCCCATC<br>AGAGTGT<br>TGCAA<br>TTTTTTG<br>ATTGTGT<br>TGCTACT<br>TTTTGGT<br>CACTATT<br>TAAAA                                                                                                                                                                                                  |
| 583.             | H3137A02-3 |                | Mus musculus<br>10 days neonate<br>cerebellum<br>cDNA, RIKEN<br>full-length<br>enriched<br>library,<br>clone:B930053<br>B19<br>product:unkno<br>wn EST, full<br>insert sequence. | H3137A02      | Mm.21657      | Chromosome 4             |                                                                                                                                                                                                                                                                                                         |
| 584.             | L0043D10-3 | A530090O15Rik  | RIKEN cDNA<br>A530090O15<br>gene                                                                                                                                                 | L0043D10      | Mm.40298      | Chromosome 15            | CTTAGGG<br>TACTAAC<br>AGAGAA1<br>GTGTATA<br>ACGTACT<br>GCTTTA<br>CATACAT<br>GCAAAAT<br>TAACTGC<br>AACCTTC<br>GTTAGTA<br>TGAGG<br>ACTTCCT<br>TACATCC<br>AGGTACA<br>GTTTACA<br>AAACTAG<br>TGAAA                                                                                                           |
| 585.             | H3087D06-3 | Etf1           | eukaryotic<br>translation<br>termination<br>factor 1                                                                                                                             | H3087D06      | Mm.3845       | Chromosome 18            |                                                                                                                                                                                                                                                                                                         |
| 586.             | C0827E01-3 |                | Mus musculus<br>15 days embryo<br>head cDNA,<br>RIKEN full-<br>length enriched<br>library,<br>clone:D930031<br>H08<br>product:unkno<br>wn EST, full<br>insert sequence.          | C0827E01      | Mm.45759      | Chromosome 10            |                                                                                                                                                                                                                                                                                                         |
| 587.             | H3053E01-3 | B130024B19Rik  | RIKEN cDNA<br>B130024B19<br>gene                                                                                                                                                 | H3053E01      | Mm.34557      | Chromosome 10            | GGAGGCA<br>AATTCCA<br>ATACAGG<br>TAAAATA<br>TAATGGG<br>GTGATT<br>AAGCGTT<br>AAGGAAA<br>CTGGAAG<br>AGGTTGT<br>CTAGCAG<br>GTCAATA<br>TTTTTTAA<br>CACTCAT<br>ACAGAGG<br>AAAGGAA<br>AGGTTTAC<br>AGTTCTC<br>AGGCATA<br>ATAGAGC<br>AGTTAGA<br>TACTCTTA<br>AAGGAGT<br>TCCTA<br>GATCACCT<br>TCCTCGAC<br>GAGATGA |
| 588.             | K0117C08-3 | BM222243       | ESTs<br>BM222243                                                                                                                                                                 | K0117C08      | Mm.221706     | Chromosome 1             |                                                                                                                                                                                                                                                                                                         |
| 589.             | H3056D11-3 | Ptgfrn         | prostaglandin<br>F2 receptor<br>negative<br>regulator                                                                                                                            | H3056D11      | Mm.24807      | Chromosome 3             |                                                                                                                                                                                                                                                                                                         |
| 590.             | C0228C02-3 | 2510004L01Rik  | RIKEN cDNA<br>2510004L01<br>gene                                                                                                                                                 | C0228C02      | Mm.24045      | Chromosome 12            |                                                                                                                                                                                                                                                                                                         |
| 591.             | H3144F09-3 | Rab711         | RAB7, member<br>RAS oncogene<br>family-like 1                                                                                                                                    | H3144F09      | Mm.34027      | Chromosome 1             |                                                                                                                                                                                                                                                                                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                          | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------|------------|----------------|---------------------------------------------------------------------------------------|---------------|---------------|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 592.             | H3052B06-3 | Abcb1b         | ATP-binding<br>cassette, sub-<br>family B<br>(MDR/TAP),<br>member 1B                  | H3052B06      | Mm.6404       | Chromosome 5             | ATGAAA<br>TTAAAAG<br>CACTTG<br>TAAAGGT<br>CCATCAA<br>AGAAGCC<br>GAGACTT<br>ATTAAAT<br>CAAAA<br>GGCCAGC<br>TGTGTAC<br>GCTCTTC<br>GGAGAAC<br>AAAACCA<br>TGGAAAT<br>CCAAGAC<br>TTTAACA<br>ATTTAAT<br>GGGTAGC<br>ATGAATG<br>GGTCCC<br>AGTGAAC<br>AAAGACA<br>AACATGT<br>TCTACTC<br>AGGAACC<br>AGAACAA<br>GATATTT<br>AGTGTCA<br>AAAAGGT<br>ATAATCT<br>TAGCGTA<br>GTAGAG<br>GTGTTAC<br>AGAAGTC<br>AGGATAA<br>AAGTTTA<br>CACAGTG<br>GAGAAG<br>TGTCCTTT<br>AATGCCA<br>GGAAGTG<br>TGCAGCT<br>GTAGAGT<br>GAGCA<br>AGAACAA<br>GGAATTT<br>CTGAAGC<br>TTTAAAG<br>TGATGTG<br>AACGCT<br>TATGGTC<br>CAAGGAA<br>AGTCACA<br>CATCTTA<br>TACTGAT<br>TAAAC<br>ATCCTCC<br>TGGTCTG<br>CATTTCCA<br>ATGTCAG<br>GTCTGCC<br>TCAGCC<br>TAAGCCC<br>TTCTGGG<br>ATCAGTT |
| 593.             | L0273B08-3 | Tgif           | TG interacting<br>factor                                                              | L0273B08      | Mm.8155       | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 594.             | K0406A08-3 | Siat4c         | sialyltransferase<br>4C (beta-<br>galactoside<br>alpha-2,3-<br>sialyltransferase<br>) | K0406A08      | Mm.2793       | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 595.             | AF075136.1 | Sap30          | sin3 associated<br>polypeptide                                                        | AF075136      | Mm.118        | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 596.             | K0644H12-3 | Prkch          | protein kinase<br>C, eta                                                              | K0644H12      | Mm.8040       | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 597.             | H3108A04-3 | Clu            | clusterin                                                                             | H3108A04      | Mm.200608     | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 598.             | H3020F06-3 | Snx10          | sorting nexin 10                                                                      | H3020F06      | Mm.29101      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 599.             | L0066C05-3 | Uxs1           | UDP-<br>glucuronate<br>decarboxylase 1                                                | L0066C05      | Mm.201248     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 600.             | L0025F08-3 | Rgs19          | regulator of G-<br>protein<br>signaling 19                                            | L0025F08      | Mm.20156      | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 601.             | H3076F06-3 | Siat4a         | sialyltransferase<br>4A (beta-<br>galactoside<br>alpha-2,3-<br>sialyltransferase<br>) | H3076F06      | Mm.248334     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 602.             | C0354G01-3 |                | Mus musculus,<br>Similar to IQ<br>motif                                               | C0354G01      | Mm.259704     | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                      | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                 |
|------------------|------------|----------------|---------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------|
|                  |            |                | containing<br>GTPase<br>activating<br>protein 2, clone<br>IMAGE:35965<br>08, mRNA,<br>partial cds |               |               |                                          | AGAGAAAC<br>GTGCAAT<br>AATGA                                      |
| 603.             | C0191H09-3 | Atp6v1a1       | ATPase, H+<br>transporting,<br>V1 subunit A,<br>isoform 1                                         | C0191H09      | Mm.29771      | No Chromosome location<br>info available | GGAAGAT<br>TTTCCAG<br>TGTATCA<br>GGACCAT<br>GTGGGGC<br>GGGAC      |
| 604.             | H3050G04-3 | Dpp7           | dipeptidylpepti<br>dase 7                                                                         | H3050G04      | Mm.21440      | Chromosome 2                             | ATGTGAT<br>AGTGGTG<br>AACTTGC<br>TATCTGA<br>CTGTCCA<br>TATGG      |
| 605.             | L0219A09-3 | Gatm           | glycine<br>amidinotransfer<br>ase (L-<br>arginine:glycine<br>amidinotransfer<br>ase)              | L0219A09      | Mm.29975      | Chromosome 2                             | AAACGAA<br>ACTTTCC<br>ATGCCTT<br>ATTCTTG<br>AACATTT<br>CTAAAC     |
| 606.             | J0821E02-3 | AU040950       | expressed<br>sequence<br>AU040950                                                                 | J0821E02      | Mm.17580      | Chromosome 13                            | AATACTC<br>TGCTGTG<br>AATTTCC<br>TACTAGA<br>GACCTCT<br>TCCTG      |
| 607.             | H3080A02-3 | Cbfb           | core binding<br>factor beta                                                                       | H3080A02      | Mm.2018       | Chromosome 8                             | GAATTAT<br>AACAAAT<br>GTTACAG<br>TGATGCT<br>TTGTGTT<br>AGCAC      |
| 608.             | C0276B08-3 | Plscr1         | phospholipid<br>scramblase 1                                                                      | C0276B08      | Mm.14627      | Chromosome 9                             | TTCTTGAC<br>TAAGGAC<br>AACTTTA<br>CCCTGAA<br>AACTGAG<br>TCACAAG   |
| 609.             | C0279E04-3 | Srd5a2l        | steroid 5 alpha-<br>reductase 2-like                                                              | C0279E04      | Mm.86611      | Chromosome 5                             | GTCACAT<br>ATAAAAA<br>GAAACTC<br>AATAATA<br>TGTACAG<br>AGACCG     |
| 610.             | K0434D04-3 | Pgd            | phosphoglucon<br>ate<br>dehydrogenase                                                             | K0434D04      | Mm.252080     | No Chromosome location<br>info available | CCCTATTC<br>ATTGATTI<br>TTCCCTTA<br>CTGTTCCC<br>TAACCCCC<br>TTTTT |
| 611.             | C0174H01-3 | Ddx21          | DEAD (Asp-<br>Glu-Ala-Asp)<br>box polypeptide<br>21                                               | C0174H01      | Mm.25264      | Chromosome 10                            | CATTGCA<br>TTTCCAAC<br>CTTTTAGA<br>ACAAAGT<br>ACCAACC<br>ATCTGC   |
| 612.             | H3085A07-3 | BG070224       | ESTs<br>BG070224                                                                                  | H3085A07      | Mm.173217     | Chromosome 17                            | TTGAGAA<br>AAAACAA<br>TCCAAAA<br>CTTTTCCT<br>GGCTATGT             |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                      | CLONE NAME | UG CLUSTER     | CHR_LOCATION PENG [A]                 | 60mer SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------|------------|---------------|--------------------------------------------------------------------------------|------------|----------------|---------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 613.       | K0208E10-3 | Mmab          | methylnmalonic aciduria (cobalamin deficiency) type B homolog (human)          | K0208E10   | Mm.105182      | Chromosome 5                          | TCGTCC<br>ACGACTC<br>TAATGTG<br>TCTCATG<br>AATTTTC<br>CCTGAAC<br>AGCAC<br>GTTGGTG<br>CTGAAAG<br>TGGAGTT<br>CAGAAAGT<br>TTGTGAT<br>TGGTTT                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 614.       | H3006F10-3 | Cops2         | COP9 (constitutive photomorphogenic) homolog, subunit 2 (Arabidopsis thaliana) | H3006F10   | Mm.3596        | Chromosome 2                          | CAGAAAA<br>AAGTCAT<br>TATGCGA<br>AGAATTA<br>ACAACTG<br>ATGTGC<br>AAATTTTC<br>TTAATTTI<br>GTCTCGA<br>AGTAACA<br>TCAACCA<br>GTCAGA<br>GGAGGAA<br>AACTGAA<br>TGTATAA<br>TAAAAAG<br>CTGATTGC<br>GGGACA<br>CAGCAGC<br>AAACACT<br>GTTAGGC<br>AGAGAAA<br>GTTAAAG<br>ATTAGAA<br>GAGAAAT<br>GTAAAAAT<br>AAAGGGA<br>ACGTGAC<br>AGGGTAG<br>GAGCTTG<br>TCAGGAA<br>TGTCATA<br>ATCTGGTA<br>TTTCTTAA<br>ATGTTGTT<br>AAGTCC<br>CAAAACA<br>ACATATTA<br>AATAAAAA<br>AAGGCGTC<br>AAATGGA<br>ACAAAAAT<br>GTAGGGA<br>TATGTCCA<br>GTTTTAGG<br>CACTTAGC<br>TAATATAC<br>TTGTAG<br>GTATACAC<br>GTAGTTAG<br>TACTGGAT<br>CTGATCAG |
| 615.       | C0108A10-3 | Nek6          | NIMA (never in mitosis gene a)-related expressed kinase 6                      | C0108A10   | Mm.143818      | Chromosome 2                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 616.       | H3028H10-3 | Ppic          | peptidylprolyl isomerase C                                                     | H3028H10   | Mm.4587        | Chromosome Multiple Mappings          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 617.       | H3121E08-3 | Ralgds        | ral guanine nucleotide dissociation stimulator                                 | H3121E08   | Mm.5236        | Chromosome 2                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 618.       | L0266H12-3 | Opa1          | optic atrophy 1 homolog (human)                                                | L0266H12   | Mm.31402       | Chromosome 16                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 619.       | K0635G02-3 | 2310046K10Rik | RIKEN cDNA 2310046K10 gene                                                     | K0635G02   | Mm.68134       | Chromosome 14                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 620.       | L0704C05-3 | 2610318G18Rik | RIKEN cDNA 2610318G18 gene                                                     | L0704C05   | Mm.180776      | Chromosome 3                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 621.       | C0303D10-3 |               | UNKNOWN C0303D10                                                               | C0303D10   | Data not found | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 622.       | K0605C04-3 | BM240648      | ESTs BM240648                                                                  | K0605C04   | Mm.265969      | Chromosome 15                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 623.       | H3071G06-3 | BG069012      | ESTs BG069012                                                                  | H3071G06   | Mm.26430       | Chromosome 4                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                            | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------|-------------|----------------|---------------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 624.             | C0600A01-3  | Coro2a         | coronin, actin<br>binding protein<br>2A                 | C0600A01      | Mm.171547     | Chromosome 4             | TTGTGTG<br>AAGTG<br>TTGTATC<br>AGGGAA/<br>GAATCAA/<br>ACGGACC<br>CTTTTCA/<br>AAACCGT<br>TGCAGCT<br>TACATTT<br>AAAAGAC<br>CCGACAC<br>TTGTAGA<br>AGGAA<br>GGCAATC<br>AATGTTG<br>CCATTCA<br>CCATGTT<br>AAATTAC<br>AGATCC<br>CCCCAA/<br>AACTGG/<br>ATTGTTT<br>CTCCTGA<br>TCTTGGA<br>CCCCCTG<br>CCAGAC/<br>ATTCTTC<br>AAATGG/<br>AAGTGA/<br>AGAATTC<br>TGTAAC<br>AGCAAA/<br>TGTATAT<br>GCTTGTC<br>AATGTCA/<br>AGGACA/<br>GAAAGA/<br>TAGAATC<br>ATTTTCT<br>CATAGTC<br>ATTGCTA<br>TAACAG/<br>ACTCAC<br>TGACGG/<br>TTGCAA/<br>AGAAAG<br>ATCTGG/<br>GCAATG/<br>TGCCTTC<br>GAAATA/<br>TGTAGCT<br>GGCTAG/<br>TGAAAA/<br>TCCAAG/<br>AGAAGG<br>ATACCA/<br>AATAAA/<br>CCAGTA/<br>AGCATC/<br>AAGATG/<br>GTCAGTC<br>CAGTGT/<br>TAGCAT/<br>TAGGTG/<br>AAAATG. |
| 625.             | NM_007679.1 | Cebpd          | CCAAT/enhanc<br>er binding<br>protein<br>(C/EBP), delta | NM_007679     | Mm.4639       | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 626.             | H3048A01-3  | Kras2          | Kirsten rat<br>sarcoma<br>oncogene 2,<br>expressed      | H3048A01      | Mm.31530      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 627.             | C0267D12-3  | Tpp2           | tripeptidyl<br>peptidase II                             | C0267D12      | Mm.28867      | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 628.             | J1012C06-3  | AU041997       | ESTs<br>AU041997                                        | J1012C06      | Mm.181004     | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 629.             | L0072F04-3  | Vav2           | Vav2 oncogene                                           | L0072F04      | Mm.179011     | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 630.             | L0836H04-3  | C030038J10Rik  | RIKEN cDNA<br>C030038J10<br>gene                        | L0836H04      | Mm.212874     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 631.             | K0614A10-3  | Sh3kbp1        | SH3-domain<br>kinase binding<br>protein 1               | K0614A10      | Mm.254904     | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 632.             | H3156B08-3  | 6620401D04Rik  | RIKEN cDNA<br>6620401D04<br>gene                        | H3156B08      | Mm.86150      | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 633.             | C0334C11-3  | B230339H12Rik  | RIKEN cDNA<br>B230339H12<br>gene                        | C0334C11      | Mm.275985     | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 634.             | H3103G05-3  | BG071839       | ESTs<br>BG071839                                        | H3103G05      | Mm.17827      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                       | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60m SEQUE                                                                                                                                                                                                                                                                           |
|------------|------------|---------------|---------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 635.       | C0205H05-3 | 1600010D10Rik | RIKEN cDNA 1600010D10 gene                                                                                                      | C0205H05   | Mm.86385   | Chromosome 3                          | GAGACTC<br>AGAATC<br>ATCCTTT<br>GTTAGTA<br>GTTTATG<br>AACTGTI<br>GAAGCTC<br>AACAGC<br>AGTGTTC<br>TGTGTAA<br>GTATTTT<br>TGGAAA/<br>GGCTGGI<br>AAGGC<br>GTCTGGC<br>TGCCCGT<br>AACCCT/<br>TTGATCA<br>AAGAAA/<br>GGTTA<br>TGTAAGA/<br>TTCTAAA<br>TGGTAAI<br>ACTCATC<br>TAAAAAI<br>CCTCG |
| 636.       | L0513G12-3 | Qk            | quaking                                                                                                                         | L0513G12   | Mm.2655    | Chromosome 17                         |                                                                                                                                                                                                                                                                                     |
| 637.       | C0100E08-3 | Pdap1         | PDGFA associated protein 1                                                                                                      | C0100E08   | Mm.188851  | Chromosome Multiple Mappings          |                                                                                                                                                                                                                                                                                     |
| 638.       | J0055B04-3 |               | transcribed sequence with strong similarity to protein pir:S12207 (M.musculus) S12207 hypothetical protein (B2 element) - mouse | J0055B04   | Mm.228682  | Chromosome 16                         | myelin basic protein                                                                                                                                                                                                                                                                |
| 639.       | J0008D10-3 | Mbp           | myelin basic protein                                                                                                            | J0008D10   | Mm.2992    | Chromosome 18                         | ACTGGAA/<br>GAATGTC<br>GCGTCGC<br>TCTGTAA<br>GGGAATC<br>TAACTT<br>TCTACTA<br>GGTTAAA/<br>ATATGAA/<br>AGAAATC<br>GAGGCTI<br>ATGCTG<br>GGACAC/<br>TTCATGT<br>TAGATTT<br>CTCGTAI<br>GCATAGC<br>GGTGG                                                                                   |
| 640.       | K0319D09-3 | Mtm1          | X-linked myotubular myopathy gene 1                                                                                             | K0319D09   | Mm.28580   | Chromosome X                          |                                                                                                                                                                                                                                                                                     |
| 641.       | C0243H05-3 | Galnt7        | UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase 7                                             | C0243H05   | Mm.62886   | Chromosome 8                          |                                                                                                                                                                                                                                                                                     |
| 642.       | L0841H10-3 | BM116846      | ESTs BM116846                                                                                                                   | L0841H10   | Mm.65363   | Chromosome 2                          | TAGATA/<br>CGTATGA<br>GAGAAA/<br>AATTAAT<br>TTCAGCA<br>GAAAGC/<br>CAATGTC<br>TGCCATC<br>AGTTTTA<br>CCTCATA<br>GTATTTA<br>TGCCC<br>CTTTGGG<br>GTTTGG<br>CCGGTTT                                                                                                                      |
| 643.       | K0334D05-3 | Ccnd1         | cyclin D1                                                                                                                       | K0334D05   | Mm.22288   | Chromosome 7                          |                                                                                                                                                                                                                                                                                     |
| 644.       | L0209B01-3 |               | L0209B01-3 NIA Mouse Newborn Ovary                                                                                              | L0209B01   |            | No Chromosome location info available |                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                           | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | cDNA Library<br>Mus musculus<br>cDNA clone<br>L0209B01 3',<br>MRNA<br>sequence<br>EST BB129550                                                                         |               |               |                                          | GGGGGGG<br>CTTTTGG<br>TTTTT                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 645.             | K0151H10-3 | BB129550       |                                                                                                                                                                        | K0151H10      | Mm.283461     | No Chromosome location<br>info available | GCCATAC<br>TATATTT<br>TGGTATC<br>GAAATCA<br>AGGAAA<br>AGTAAA<br>TGGTGTT<br>TTACAGT<br>CATCACA<br>ATCTAAA<br>CTTCGTT<br>CCAGC                                                                                                                                                                                                                                                                                                                                                    |
| 646.             | L0505B11-3 | Ammecr1        | Alport<br>syndrome,<br>mental<br>retardation,<br>midface<br>hypoplasia and<br>elliptocytosis<br>chromosomal<br>region gene 1<br>homolog<br>(human)<br>ESTs<br>BM120800 | L0505B11      | Mm.143724     | Chromosome X                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 647.             | L0944C06-3 | BM120800       |                                                                                                                                                                        | L0944C06      | Mm.217092     | Chromosome 3: not placed                 | TATTTGG<br>AAAGAA<br>GTTGAAA<br>TCATCCA<br>CATGCAI<br>TAACAC<br>CGAGGAC<br>TAGGGAC<br>CATGGAC<br>ATAAGAA<br>CTTGGGC<br>AAAGAGC<br>TGGTGAC<br>ATTACGT<br>ATCTCTG<br>TGTGATA<br>CGATAAA<br>TAAGAG<br>ACCCTTT<br>AAATAGT<br>AAAACGT<br>TGTTTAG<br>ATATAAA<br>ATGCAGC<br>GTTGGAC<br>ATACAAC<br>CATTGAA<br>GAACAAC<br>TTATTGT<br>TAACAG<br>TTCCTAC<br>TGTGTTT<br>AGGATTA<br>AGTAGCG<br>TGTA CTG<br>GAAAC<br>TAGATAA<br>GACTATT<br>ATTTTAG<br>AGAAAGT<br>CATGCGT<br>CTACCT<br>GGGGGGG |
| 648.             | J0027C07-3 | Mrps25         | mitochondrial<br>ribosomal<br>protein S25                                                                                                                              | J0027C07      | Mm.87062      | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 649.             | L0855B04-3 | Wdr26          | WD repeat<br>domain 26                                                                                                                                                 | L0855B04      | Mm.21126      | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 650.             | H3060H05-3 |                | Mus musculus<br>cDNA clone<br>MGC:28609<br>IMAGE:42185<br>51, complete<br>cds                                                                                          | H3060H05      | Mm.11778      | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 651.             | K0330G09-3 | 5830461H18Rik  | RIKEN cDNA<br>5830461H18<br>gene                                                                                                                                       | K0330G09      | Mm.261448     | Chromosome 14                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 652.             | L0803E07-3 | Dpysl4         | dihydropyrimid<br>inase-like 4                                                                                                                                         | L0803E07      | Mm.250414     | Chromosome 7                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 653.             | L0283B01-3 | Ivns1abp       | influenza virus<br>NS1A binding<br>protein                                                                                                                             | L0283B01      | Mm.33764      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 654.             | L0065G02-3 | 6530401D17Rik  | RIKEN cDNA                                                                                                                                                             | L0065G02      | Mm.27579      | Chromosome X                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 6530401D17<br>gene                                                                                                                                                                         |               |               |                                          | TTAATAT<br>TGTTAGA<br>ATAAGTG<br>AAATAAA<br>ACTAAAG<br>AAAAGAG<br>CTGTCCT<br>CTCAACT<br>AGTACTC<br>TAAGATG<br>ATTGTC                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 655.             | C0949A06-3 |                | Mus musculus<br>0 day neonate<br>skin cDNA,<br>RIKEN full-<br>length enriched<br>library,<br>clone:4632424<br>N07<br>product:unkno<br>wn EST, full<br>insert sequence.<br>ESTs<br>BG071548 | C0949A06      | Mm.71633      | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 656.             | H3100C11-3 | BG071548       |                                                                                                                                                                                            | H3100C11      | Mm.173983     | Chromosome Un: not<br>placed             | CAAATGT<br>AGAAACA<br>TCATGAA<br>CTTGAAA<br>CTTCTTA<br>AGCTCC<br>AACATAA<br>AAATATA<br>GGAATAT<br>AATTAAA<br>ATGTTTT<br>TTAGT<br>GACTATT<br>AGATTAG<br>GTCATGT<br>CTCGTCA<br>AGCCAAA<br>TCTGTG<br>ACAAACA<br>GAAAAAA<br>AGTAGGA<br>GGAGAAA<br>CTCACAG<br>GAATGTT<br>AATTTCAC<br>GGCTTAC<br>ATGTAAAC<br>TCCTGTAA<br>ACTCATGT<br>ACATC<br>TATACCA<br>GAAAACG<br>AATCTCA<br>AAGTAAG<br>GGTTTTGT<br>CCCTGC<br>TAGCCATT<br>GAGATGT<br>TCAAAGTC<br>TGATGATC<br>TTGCACTT<br>AATCA<br>GCTCAGCT<br>GCTAGACT<br>ACCAGGT<br>CAGAAAG<br>GAGAAAC<br>ACTCAGCA |
| 657.             | C0142H08-3 | 3110020O18Rik  | RIKEN cDNA<br>3110020O18<br>gene                                                                                                                                                           | C0142H08      | Mm.117055     | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 658.             | L0945G09-3 | Bcl2l11        | BCL2-like 11<br>(apoptosis<br>facilitator)                                                                                                                                                 | L0945G09      | Mm.141083     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 659.             | L0848H06-3 | E130318E12Rik  | RIKEN cDNA<br>E130318E12<br>gene                                                                                                                                                           | L0848H06      | Mm.198119     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 660.             | K0617B02-3 | Bmp2k          | BMP2<br>inducible<br>kinase                                                                                                                                                                | K0617B02      | Mm.6156       | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 661.             | C0203D07-3 | Pfk1           | PFTAIRE<br>protein kinase 1                                                                                                                                                                | C0203D07      | Mm.6456       | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 662.             | L0267A02-3 | 2210409B22Rik  | RIKEN cDNA<br>2210409B22<br>gene                                                                                                                                                           | L0267A02      | Mm.30015      | Chromosome 4                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 663.             | J0086F05-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>sp:P00722 (E.<br>coli)<br>BGAL_ECOLI                                                                               | J0086F05      | Mm.31079      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------|------------|----------------|-----------------------------------------------------------------|---------------|---------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 664.             | C0606A03-3 | Rps23          | Beta-<br>galactosidase<br>(Lactase)<br>ribosomal<br>protein S23 | C0606A03      | Mm.295618     | Chromosome X             | TATCACT<br>TATTGAA<br>TGTATGT<br>TGGGAG/<br>ACTTTCT<br>TAAGGT<br>ACTGCTG<br>AAACAA/<br>ACTACAT<br>CAATAGT<br>TACCATG<br>TGGCG<br>GAAGGA/<br>CAAACAC<br>GAACTTC<br>CTTTCAG<br>AAAACA/<br>TTGTCCC<br>AGAAAA/<br>TAAACTC<br>TTAGTAT<br>ACGAGCA<br>AGTGGTG<br>AAGCTCC<br>AAAGGAA/<br>AAGAGTG<br>ATTTGGA/<br>GAAAGAT<br>CAGTTTA/<br>AAAGAC<br>GAAATGG<br>TGAGGCT<br>AAATGAA<br>GGCTAGT<br>CAAAGAT<br>GTATCC<br>ACTATTTC<br>TCAATAG<br>GCAAAAG<br>ACTAATT<br>TGTATAT/<br>AGTGTA<br>TCCTCTA/<br>TGTGTCT/<br>TACATGA/<br>CATTGGT/<br>TCAAACA<br>GGGTG<br>TTGGAAA/<br>AAGTAAC<br>AGACGGC<br>ATTCTTA/<br>CCGAAAA<br>ACCCCAA<br>GTGTGAT/<br>TTTTCATC<br>CTAGAGC<br>GACAAAG<br>TTACTCTT<br>TCGCAA<br>GGCTTTAC<br>AAACTTC/<br>TTCAAAG/ |
| 665.             | L0902D02-3 | Ncoa6ip        | nuclear receptor<br>coactivator 6<br>interacting<br>protein     | L0902D02      | Mm.171323     | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 666.             | H3060C12-3 | BG067974       | ESTs<br>BG067974                                                | H3060C12      | Mm.173106     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 667.             | C0611E01-3 | Tor3a          | torsin family 3,<br>member A                                    | C0611E01      | Mm.206737     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 668.             | U54984.1   | Mmp14          | matrix<br>metalloproteinase 14<br>(membrane-<br>inserted)       | U54984        | Mm.19945      | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 669.             | H3089F08-3 | 0610013E23Rik  | RIKEN cDNA<br>0610013E23<br>gene                                | H3089F08      | Mm.182061     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 670.             | K0633C04-3 | Ebi2           | Epstein-Barr<br>virus induced<br>gene 2                         | K0633C04      | Mm.265618     | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 671.             | J0943E09-3 | Nup62          | nucleoporin 62                                                  | J0943E09      | Mm.22687      | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 672.             | L0267D03-3 | Dcn            | decorin                                                         | L0267D03      | Mm.56769      | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 673.             | L0250B09-3 | 1110031E24Rik  | RIKEN cDNA<br>1110031E24<br>gene                                | L0250B09      | Mm.34356      | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 674.             | L0915B12-3 | Etv3           | ets variant gene<br>3                                           | L0915B12      | Mm.34510      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                                  | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|------------------|-------------|----------------|---------------------------------------------------------------|---------------|---------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 675.             | NM_009403.1 | Tnfsf8         | tumor necrosis<br>factor (ligand)<br>superfamily,<br>member 8 | NM_009403     | Mm.4664       | Chromosome 4             | CTTCTAA<br>TTCCTTC<br>AAAAA<br>AAAGTAC<br>ATGAGAT<br>ATTTCCC<br>ATTTTCT<br>CTCAGAA<br>GAGACTC<br>AGTCCTC<br>TGTTTCC<br>TTTCCTT<br>TGAAGGC<br>TTGGATC<br>CTTAC<br>AAGAAT/<br>CACTTGA<br>ATACTGT<br>GGAAATC<br>ACTGTTT<br>AAAACCT<br>GTTAGAT<br>TTGAAGC<br>AATAACT<br>CTAATAC<br>GAAAAC<br>ACTAAG<br>AGCAGAT<br>GACTTCT<br>TACACAC<br>GCTAACT<br>TGTATGA<br>TACAG<br>TGTCTAT<br>GAAGTA/<br>CCTGAA/<br>ATAAGGC<br>ACAAAC/<br>TTACTT<br>GGGAAG/<br>AGAATT<br>GAAGATC<br>GGTTTTT<br>TTTTTTT<br>TTTACA<br>CTTGAAG<br>AGTATA<br>TAGGCA<br>GAGAAA<br>TTTGATC<br>CTGGTTA<br>TCCTTCA<br>GATATCT<br>CAGAGA/<br>AAAATA/<br>GCATGG<br>AAATGAC<br>TATGGAA<br>GAAATA/<br>CATCTGT<br>AAGAACC<br>GATGGAA<br>ATACCGC<br>AGGTCAA<br>AAGTTTT<br>GTTTAAT |
| 676.             | C0308F04-3  | 2700064H14Rik  | RIKEN cDNA<br>2700064H14<br>gene                              | C0308F04      | Mm.24730      | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 677.             | C0288G12-3  | 6030400A10Rik  | RIKEN cDNA<br>6030400A10<br>gene                              | C0288G12      | Mm.159840     | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 678.             | H3005A11-3  | Fancd2         | Fanconi<br>anemia,<br>complementatio<br>n group D2            | H3005A11      | Mm.291487     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 679.             | H3121H07-3  | 2810405I11Rik  | RIKEN cDNA<br>2810405I11<br>gene                              | H3121H07      | Mm.73777      | Chromosome 18            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 680.             | K0124A06-3  | BM222608       | ESTs<br>BM222608                                              | K0124A06      | Mm.221709     | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 681.             | NM_010835.1 | Msx1           | homeo box,<br>msh-like 1                                      | NM_010835     | Mm.259122     | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 682.             | K0134C07-3  | Falz           | fetal Alzheimer<br>antigen                                    | K0134C07      | Mm.218530     | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 683.             | K0424H02-3  | Pfkl           | phosphofructok<br>inase, platelet                             | K0424H02      | Mm.108076     | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 684.             | H3153G06-3  | 8030446C20Rik  | RIKEN cDNA<br>8030446C20<br>gene                              | H3153G06      | Mm.204920     | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 685.             | H3071C09-3  | BG068971       | ESTs<br>BG068971                                              | H3071C09      | Mm.162073     | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                                                            | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A]                 | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------|-------------|----------------|-----------------------------------------------------------------------------------------|---------------|----------------|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 686.             | L0243B07-3  |                | Possibly<br>intronic in<br>U008124-<br>L0243B07                                         | L0243B07      | Data not found | No Chromosome location<br>info available | AGTTAGC<br>AAGACTT<br>CACGG<br>AATGCTT<br>TTGAGTC<br>TGTTTAC<br>CCTATGA<br>GCATTTT<br>ACAAC                                                                                                                                                                                                                                                                                                                                                                                                         |
| 687.             | C0143D11-3  | Ii             | Ia-associated<br>invariant chain                                                        | C0143D11      | Mm.248267      | Chromosome 18                            | TAAAGGC<br>CCCCATT<br>ACCCATT<br>GTCTTGA<br>GGGGCTC<br>ATAAAG<br>CCCCTTT<br>AACTGGC<br>AAATCCT<br>AGAAAGC<br>ATTTAGA<br>TGCCCC                                                                                                                                                                                                                                                                                                                                                                      |
| 688.             | L0512A02-3  | Snx5           | sorting nexin 5                                                                         | L0512A02      | Mm.20847       | No Chromosome location<br>info available | GTCAAGT<br>GGTTTCC<br>CATCAGG<br>AATGGAT<br>TAAAGAC<br>GGGCGTT<br>GAAAGCC<br>AGCGAA<br>TCTCGTG<br>GTTGAAT<br>TCCAAAC<br>AAATAT<br>GAAATAT<br>ACTAAGA<br>GCCCAAA<br>ACTGGAT<br>TTATCCA<br>CTTAGTT<br>GTATACA<br>TATTTT<br>TAAGGCC<br>CTTCTGA<br>CTTGGA<br>CAGAG<br>GGATGAA<br>GAAGATT<br>GCAGGTC<br>AAACCTG<br>TCTAGTA<br>TCACTCT<br>TTCAATC<br>AAGTAGA<br>AGTTCTT<br>ATCTGTT<br>TTCAGAA<br>CTCAG<br>AAATTTT<br>AAGCTAT<br>TCTGACT<br>ATTTTGTC<br>CCATTTAC<br>AAACT<br>AGAATCT<br>ACTAAAG<br>GTATAGA |
| 689.             | K0112C06-3  | Atp8a1         | ATPase,<br>aminophospholipid<br>transporter<br>(APLT), class I,<br>type 8A,<br>member 1 | K0112C06      | Mm.200366      | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 690.             | H3053A01-3  | Tnfsf13b       | tumor necrosis<br>factor (ligand)<br>superfamily,<br>member 13b                         | H3053A01      | Mm.28835       | Chromosome 8                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 691.             | C0668F08-3  | Atp6ap2        | ATPase, H <sup>+</sup><br>transporting,<br>lysosomal<br>accessory<br>protein 2          | C0668F08      | Mm.25148       | Chromosome X                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 692.             | K0417E05-3  | Osmr           | oncostatin M<br>receptor                                                                | K0417E05      | Mm.10760       | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 693.             | NM_010872.1 | Birc1b         | baculoviral IAP<br>repeat-<br>containing 1b                                             | NM_010872     | Mm.89961       | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 694.             | L0262G06-3  | Cfh            | complement<br>component<br>factor h                                                     | L0262G06      | Mm.8655        | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 695.             | J0249F06-3  | 2210023K21Rik  | RIKEN cDNA<br>2210023K21<br>gene                                                        | J0249F06      | Mm.28890       | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 696.             | C0170A02-3  | Serpinb9       | serine (or<br>cysteine)<br>proteinase                                                   | C0170A02      | Mm.3368        | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------|------------|----------------|-----------------------------------------------------------------|---------------|---------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | inhibitor, clade<br>B, member 9                                 |               |               |                          | ACTGTTC<br>GTTTTC<br>AGGCC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 697.             | H3076C12-3 | FacI4          | fatty acid-<br>Coenzyme A<br>ligase, long<br>chain 4            | H3076C12      | Mm.143689     | Chromosome X             | ATCTTTG<br>TATTTTC<br>TAGCAT/<br>AAATGT/<br>CAGTGA/<br>CTGAGA<br>GGGTTA/<br>CACTGA/<br>AGAAGT/<br>AAAACT/<br>AATGTA/<br>GGAAAG<br>TACTTGT<br>CAAGCT/<br>AAGTTA/<br>AGAGAA<br>CGAACT/<br>GAGCAA/<br>TAAATA/<br>TTCCCA/<br>CCACTG/<br>AATGGA/<br>CTGTCC/<br>TTCAAT<br>AAATAG/<br>TTTAAG/<br>AGGAAG<br>ATTCCG/<br>TCACAG/<br>TCAAGG<br>TGAATC/<br>GCAACT/<br>TCTCTG/<br>CTACCT/<br>CTCTTG/<br>AGCTG<br>TGGCAA.<br>TAGATG.<br>AATGTT/<br>TAAATC/<br>ACTCAT/<br>ACTTTG/<br>ACCACG/<br>ATGACC.<br>CAGGAT.<br>AGTTTT/<br>AAATTT/<br>GCCTGG<br>GACATC/<br>CTCTCT/<br>CAGTAG/<br>TCATCG/<br>GCCATT/<br>ATGGG<br>TCTGTG/<br>CTCATG/<br>GTCTGA.<br>CACCTC.<br>AGATGT/<br>GAATT<br>TTCCAG/<br>ATGTCT/<br>TTTCAAC |
| 698.             | H3155C07-3 | 1810036L03Rik  | RIKEN cDNA<br>1810036L03<br>gene                                | H3155C07      | Mm.27385      | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 699.             | K0331C04-3 | Sdcccag8       | serologically<br>defined colon<br>cancer antigen<br>8           | K0331C04      | Mm.171399     | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 700.             | J0538B04-3 | Laptm5         | lysosomal-<br>associated<br>protein<br>transmembrane<br>5       | J0538B04      | Mm.4554       | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 701.             | H3014E07-3 | 1810029G24Rik  | RIKEN cDNA<br>1810029G24<br>gene                                | H3014E07      | Mm.27800      | Chromosome 18            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 702.             | K0515H12-3 | 2900064A13Rik  | RIKEN cDNA<br>2900064A13<br>gene                                | K0515H12      | Mm.268027     | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 703.             | H3159D10-3 | BG076403       | ESTs<br>BG076403                                                | H3159D10      | Mm.103300     | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 704.             | K0127F02-3 | Prg            | proteoglycan,<br>secretory<br>granule                           | K0127F02      | Mm.22194      | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 705.             | L0919B08-3 | Bnip31         | BCL2/adenovir<br>us E1B 19kDa-<br>interacting<br>protein 3-like | L0919B08      | Mm.29820      | Chromosome 14            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 706.             | J0904A09-3 | 1110060F11Rik  | RIKEN cDNA<br>1110060F11<br>gene                                | J0904A09      | Mm.4859       | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 707.             | L0270B06-3 | D11Ertd759e    | DNA segment,<br>Chr 11,<br>ERATO Doi                            | L0270B06      | Mm.30111      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                      | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------|------------|----------------|---------------------------------------------------|---------------|---------------|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 759, expressed                                    |               |               |                                 | GATGTG<br>GTAAGC<br>TCCGA<br>AACCAT<br>AAATGC<br>AGATAA<br>GAGATTC<br>AATGCC<br>TTAGCT<br>GTGAATC<br>GTTTACT<br>GTAAGA<br>AGAAAA<br>AACTACA<br>CTATGAC<br>TTCACAA<br>GACACA<br>TGGAAG<br>AACTGAC<br>AAGTCTT<br>CTGAG<br>GAAGAT<br>GATGTAT<br>GTGGCGT<br>TCCAGTA<br>CTGTCAT<br>CTCCA<br>AGAATGA<br>AGAATGC<br>AAACGTA<br>TTTGAAG<br>TCGTTGA<br>CTATTTG<br>TCGACAA<br>GTAATCC<br>AATGGAC<br>AAAACCT<br>GCACTTC<br>ATATACA<br>TATATGC<br>TCATAGA<br>CTGCAAT<br>ACTTAGC<br>TAAGCAT<br>ATAGAC<br>CGTCATA<br>CTATTTG<br>CAAGAGC<br>GACTACA<br>GAAGATA<br>TGCATAG<br>CTCAGAT<br>TCTTTAG<br>AGCTGGT<br>AAATGGG<br>GTAAAAAC<br>GAAGC<br>AATGAAA<br>GCGTCTA<br>TTGAAAG<br>TGTTAAC<br>TTGAATG<br>TTCCC<br>AATCTTC<br>AGACATT<br>ATTTGAA |
| 708.             | K0230D06-3 | Eaf1           | ELL associated factor 1                           | K0230D06      | Mm.37770      | Chromosome 14                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 709.             | K0611A03-3 | A1447904       | expressed sequence<br>A1447904                    | K0611A03      | Mm.447        | Chromosome 1                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 710.             | H3155A07-3 | BG076050       | ESTs<br>BG076050                                  | H3155A07      | Mm.182857     | Chromosome 5                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 711.             | H3028H11-3 | Ctsh           | cathepsin H                                       | H3028H11      | Mm.2277       | Chromosome Multiple<br>Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 712.             | L0001D12-3 | 4833422F06Rik  | RIKEN cDNA<br>4833422F06<br>gene                  | L0001D12      | Mm.27436      | Chromosome 15                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 713.             | L0951G01-3 | BG061831       | ESTs<br>BG061831                                  | L0951G01      | Mm.133824     | Chromosome 10                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 714.             | H3035G02-3 | A1314180       | expressed sequence<br>A1314180                    | H3035G02      | Mm.27829      | Chromosome 4                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 715.             | C0925G02-3 | Fer1l3         | fer-1-like 3,<br>myoferlin (C.<br>elegans)        | C0925G02      | Mm.34674      | Chromosome 19                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 716.             | C0103H10-3 | Il17r          | interleukin 17<br>receptor                        | C0103H10      | Mm.4481       | Chromosome 6                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 717.             | H3129F05-3 | Mrpl16         | mitochondrial<br>ribosomal<br>protein L16         | H3129F05      | Mm.203928     | Chromosome 19                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 718.             | L0942B12-3 |                | Mus musculus<br>12 days embryo<br>spinal ganglion | L0942B12      | Mm.214553     | Chromosome 15                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                           |
|------------------|------------|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | cDNA, RIKEN<br>full-length<br>enriched<br>library,<br>clone:D130046<br>C24<br>product:unkno<br>wn EST, full<br>insert sequence.<br>phospholipase<br>C, gamma 2                  |               |               |                          | CCTGAAA<br>TTAGAAA<br>CAGGC                                                                                                                                                                                                                                                                                               |
| 719.             | L0009B09-3 | Plcg2          |                                                                                                                                                                                 | L0009B09      | Mm.22370      | Chromosome 8             | TACCCCA<br>AGGCATC<br>CCGGGTT<br>TCAGTCC<br>GAAGAA1<br>TACAGT<br>TTTTTTC1<br>CCAATGT<br>TTGTAAG<br>GTAAATA<br>ATTTTGA.<br>AACA<br>CACACCC<br>ATGTTCC.<br>GCTCCAG<br>AGATCTT<br>CTCATGA.<br>TGACA<br>CCCAGGT.<br>CTAAGCA<br>AGGTTTG.<br>CATTTAC<br>TTCAAAT.<br>GACGG                                                       |
| 720.             | C0665B08-3 | Sh3bp1         | SH3-domain<br>binding protein<br>1                                                                                                                                              | C0665B08      | Mm.4462       | Chromosome 15            |                                                                                                                                                                                                                                                                                                                           |
| 721.             | H3102F04-3 | Rgs10          | regulator of G-<br>protein<br>signalling 10                                                                                                                                     | H3102F04      | Mm.18635      | Chromosome 7             |                                                                                                                                                                                                                                                                                                                           |
| 722.             | K0547F06-3 |                | transcribed<br>sequence with<br>moderate<br>similarity to<br>protein<br>sp:P00722 (E.<br>coli)<br>BGAL_ECOLI<br>Beta-<br>galactosidase<br>(Lactase)<br>galactosidase,<br>beta 1 | K0547F06      | Mm.162929     | Chromosome 19            |                                                                                                                                                                                                                                                                                                                           |
| 723.             | H3087C07-3 | Glb1           |                                                                                                                                                                                 | H3087C07      | Mm.255070     | Chromosome 9             | GGAGCAA<br>TTGAATA.<br>CCTTTATC<br>ATTTGAA.<br>TCACGTC.<br>TTCTGC<br>TGGAATA.<br>AAGAATC<br>GTAGAAA<br>AGACTTGC<br>ATAGGGT<br>TAAGGC<br>ACCACAG<br>TCAGCAT<br>AGATTTC<br>ATGATCC<br>TTGTCTTG<br>TAGGG<br>AGGGTCAC<br>CCGAATC<br>GGACACAC<br>ACAAGGA<br>TAATCCAA<br>GATGTAT<br>AGTGGAG<br>CAGTCTGC<br>TTCAGGAT<br>GTGAATAA |
| 724.             | J0437D05-3 | AU023716       | ESTs<br>AU023716                                                                                                                                                                | J0437D05      | Mm.173654     | Chromosome X             |                                                                                                                                                                                                                                                                                                                           |
| 725.             | H3156A09-3 | Pex12          | peroxisomal<br>biogenesis<br>factor 12                                                                                                                                          | H3156A09      | Mm.30664      | Chromosome 11            |                                                                                                                                                                                                                                                                                                                           |
| 726.             | G0108H12-3 | Ly6e           | lymphocyte<br>antigen 6<br>complex, locus<br>E                                                                                                                                  | G0108H12      | Mm.788        | Chromosome 15            |                                                                                                                                                                                                                                                                                                                           |
| 727.             | H3098D12-5 | Map2k1         | mitogen<br>activated<br>protein kinase<br>kinase 1                                                                                                                              | H3098D12      | Mm.248907     | Chromosome 9             |                                                                                                                                                                                                                                                                                                                           |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                                                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60me<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 728.             | C0637C02-3 | Zmpste24       | zinc<br>metalloproteinase, STE24<br>homolog (S.<br>cerevisiae)                                                                                                                                                             | C0637C02      | Mm.34399      | Chromosome 4                    | CTTAATA<br>ACCCT<br>TTTGGGC<br>AAAAAC/<br>TCAGTTT<br>CAAGTGA<br>CTTAAAA<br>CCCATG<br>AAAGGA/<br>AAAGTGC<br>GAAAGT/<br>TCTGCTT<br>GCATGTG<br>TGGTGCC<br>TTCACTC<br>ACTGTGA<br>CAGTGGC<br>TGGAAA/<br>CAGAGA/<br>AACTGTC<br>CACCATC<br>CAGAATA/<br>ATGAAA/<br>ATGCAA/<br>GTAAGCT<br>CTCAT<br>TTGTGGA<br>GAAATA/<br>ATAATTG<br>CCTCTAG<br>TGGATCT<br>ATGTTG<br>ACCAGA/<br>ACAGTCT<br>TTCAGCC<br>GGACTCC<br>CTGAGA/<br>GTAACA/<br>GATACTC<br>CTTTGAA<br>AAGAAC/<br>GCTAAA/<br>TGAAGCT<br>GGTGGC |
| 729.             | H3119B06-3 | Atp1b3         | ATPase,<br>Na+/K+-<br>transporting,<br>beta 3<br>polypeptide                                                                                                                                                               | H3119B06      | Mm.424        | Chromosome Multiple<br>Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 730.             | C0176B06-3 | Ubl1           | ubiquitin-like 1                                                                                                                                                                                                           | C0176B06      | Mm.259278     | Chromosome Multiple<br>Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 731.             | C0626D04-3 | 9130404D14Rik  | RIKEN cDNA<br>9130404D14<br>gene                                                                                                                                                                                           | C0626D04      | Mm.219676     | Chromosome 2                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 732.             | H3155E07-3 | Dock4          | dedicator of<br>cytokinesis 4                                                                                                                                                                                              | H3155E07      | Mm.145306     | Chromosome 12                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 733.             | C0106A05-3 | H2-Eb1         | histocompatibili<br>ty 2, class II<br>antigen E beta                                                                                                                                                                       | C0106A05      | Mm.22564      | Chromosome 17                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 734.             | H3037B09-3 |                | Mus musculus<br>12 days embryo<br>spinal cord<br>cDNA, RIKEN<br>full-length<br>enriched<br>library,<br>clone:C530028<br>D16<br>product:231000<br>8H09RIK<br>PROTEIN<br>homolog [Mus<br>musculus], full<br>insert sequence. | H3037B09      | Mm.274876     | Chromosome 7                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 735.             | H3003B09-3 | F730017H24Rik  | RIKEN cDNA<br>F730017H24<br>gene                                                                                                                                                                                           | H3003B09      | Mm.205421     | Chromosome 14                   | CCATTTC<br>TCACTGC<br>TTAGTGC<br>GAGAAA<br>TTTTTAA<br>ATCTTG<br>GGCAAC/<br>AAGTGT/<br>TTCTAAC<br>AAACTG/<br>AACTTG/<br>ATACTG<br>CAGAAG                                                                                                                                                                                                                                                                                                                                                    |
| 736.             | C0909E10-3 | Pign           | phosphatidylinositol<br>glycan,<br>class N                                                                                                                                                                                 | C0909E10      | Mm.268911     | Chromosome 1                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 737.             | H3045G01-3 | BG066588       | ESTs                                                                                                                                                                                                                       | H3045G01      | Mm.269064     | Chromosome 14                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                         | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | BG066588                                                                                                                                                             |               |               |                          | TCTGAAA<br>TAGTTGT<br>ACTCTAA<br>GATCCAT<br>GAAAAG<br>TATCGTA<br>GCACCTA<br>TAAGTGG<br>ATGCTCT<br>ACACTCA<br>AGCTGGG                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 738.             | H3006E10-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>sp:Q9H321<br>(H.sapiens)<br>VCXC_HUMA<br>N VCX-C<br>protein<br>(Variably<br>charged protein<br>X-C) | H3006E10      | Mm.218665     | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 739.             | H3098H09-3 | 2310016E02Rik  | RIKEN cDNA<br>2310016E02<br>gene                                                                                                                                     | H3098H09      | Mm.21450      | Chromosome 5             | TGTTTTGT<br>TAAATCA<br>CACTCAC<br>TCTCCCA<br>CTGATAA<br>TTTAC<br>AGCCACT<br>CTCTAAA<br>AATTTCA<br>CTTGAGT<br>TCCTCTA<br>GTTTA<br>GCTTTGT<br>ATGGTCA<br>CCCAAAC<br>GAGCCTT<br>ATGTGTT<br>GACCT<br>CCTTAGA<br>TGGTAAT<br>TTTAGGT<br>GTACTAT<br>CGCCATT<br>AACCC<br>TAAATG<br>CTTTTGG<br>AAAGATG<br>ACGTAGA<br>AGTGCTA<br>ACGTTTC<br>GCAGTTA<br>TCTTTGG<br>TCACAAC<br>GTGACAT<br>TCCTTTG<br>AAGCA<br>TGCTTAG<br>ACATAGA<br>GAAGCAA<br>GGATGCC<br>CACTGAG<br>AGGTTTC<br>GGTTTTC<br>CACGTAC<br>ATGCCTC<br>TTGTGAA<br>TGACTTT<br>AACCC<br>GTTCACTC |
| 740.             | J0540D09-3 | Adam9          | a disintegrin<br>and<br>metalloproteinase<br>domain 9<br>(meltrin<br>gamma)                                                                                          | J0540D09      | Mm.28908      | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 741.             | L0208C06-3 | Pknox1         | Pbx/knotted 1<br>homeobox                                                                                                                                            | L0208C06      | Mm.259295     | Chromosome 17            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 742.             | H3154G05-3 | Napg           | N-<br>ethylmaleimide<br>sensitive fusion<br>protein<br>attachment<br>protein gamma                                                                                   | H3154G05      | Mm.154623     | Chromosome 18            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 743.             | L0854E11-3 | 1500032M01Rik  | RIKEN cDNA<br>1500032M01<br>gene                                                                                                                                     | L0854E11      | Mm.29628      | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 744.             | H3014C06-3 | B2m            | beta-2<br>microglobulin                                                                                                                                              | H3014C06      | Mm.163        | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 745.             | K0538G12-3 | Ccr2           | chemokine (C-<br>C) receptor 2                                                                                                                                       | K0538G12      | Mm.6272       | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 746.             | J0819C09-3 | C030002B11Rik  | RIKEN cDNA<br>C030002B11<br>gene                                                                                                                                     | J0819C09      | Mm.70065      | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 747.             | C0175B11-3 | Hist1h2bc      | histone 1, H2bc                                                                                                                                                      | C0175B11      | Mm.21579      | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                    | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                               |
|------------------|------------|----------------|---------------------------------------------------------------------------------|---------------|---------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| 748.             | H3009B11-3 | Nufip1         | nuclear fragile<br>X mental<br>retardation<br>protein<br>interacting<br>protein | H3009B11      | Mm.21138      | Chromosome 14            | AAATTTG<br>AAGAAAC<br>CACAGAC<br>GAAAATC<br>ATACTTG<br>AAAGAC<br>TGGACTT<br>CTGATTC<br>AAAACTC<br>AAGTGTA<br>TCTCCC           |
| 749.             | H3135D02-3 | Lamp2          | lysosomal<br>membrane<br>glycoprotein 2                                         | H3135D02      | Mm.486        | Chromosome X             | CTGGTGT<br>TATTTTC<br>CTTTAGA<br>GTATAAC<br>CTGGTCC<br>AAGTAC<br>TAAAGGT<br>GTGTCTC<br>CCCCAGC<br>GGAGATT<br>CAACTAT<br>GGGGT |
| 750.             | K0540G08-3 | 1200013B08Rik  | RIKEN cDNA<br>1200013B08<br>gene                                                | K0540G08      | Mm.247440     | Chromosome X             | CTGAATT<br>TCACTTG<br>TTCTCAT<br>ACCTCCA<br>CAACAA<br>ATGTCT<br>TGTGCTT<br>AAAATGC<br>ATAATTC<br>TTAGAGC<br>TATCAAC<br>CCTTAC |
| 751.             | H3089H05-3 | Ln timer       | ligand of numb-<br>protein X 2                                                  | H3089H05      | Mm.34462      | Chromosome 5             | TCTGTGA<br>TTGTAGA<br>CCGTAA<br>ATCCAG<br>TAGCAG<br>GGAAAG<br>CTTACA<br>TCCTAA<br>TGGGCC<br>TTCTTT<br>GGTTGA<br>ATGAA         |
| 752.             | J0203A08-3 | C85149         | ESTs C85149                                                                     | J0203A08      | Mm.154684     | Chromosome 2             | CTGTTTA<br>ATGAAA<br>GAAGCT<br>GAAGAC<br>AGACGA<br>CATTGTA<br>TGAATA<br>GGGCCA<br>ATATAA<br>ATCCAG<br>ATGGCT<br>TGTGC         |
| 753.             | H3119F01-3 | Mcfid2         | multiple<br>coagulation<br>factor<br>deficiency 2                               | H3119F01      | Mm.30251      | Chromosome 17            | GGGGGA<br>CTATATC<br>TCGTTT<br>TGACTT<br>GATAGT<br>AACTTC<br>AAACTTC                                                          |
| 754.             | H3134C05-3 | Mglap          | matrix gamma-<br>carboxyglutamate (gla) protein                                 | H3134C05      | Mm.243085     | Chromosome 6             |                                                                                                                               |
| 755.             | C0147D11-3 | B230215M10Rik  | RIKEN cDNA<br>B230215M10<br>gene                                                | C0147D11      | Mm.41525      | Chromosome 10            |                                                                                                                               |
| 756.             | C0949H10-3 | Sulf1          | sulfatase 1                                                                     | C0949H10      | Mm.45563      | Chromosome 1             |                                                                                                                               |
| 757.             | K0114E04-3 | BM222075       | ESTs<br>BM222075                                                                | K0114E04      | Mm.221705     | Chromosome 19            |                                                                                                                               |
| 758.             | H3012C03-3 | Cappa1         | capping protein                                                                 | H3012C03      | Mm.19142      | Chromosome Multiple      |                                                                                                                               |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                       | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------|------------|----------------|--------------------------------------------------------------------|---------------|----------------|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | alpha 1                                                            |               |                | Mappings                 | ACACAG/<br>GAAGGA/<br>TAGGTAT<br>GCTTTAT<br>TCTGGCA<br>AATAAGC<br>AAGAATT<br>TTGGAAA<br>ATACACC<br>TTAGGCA<br>CAAGGC<br>TCCCACT<br>ACAGATC<br>TCTTGTG<br>GGTGCCA<br>CTGGTAC<br>GGCCT                                                                                                                                                                                                                                                                                                                                                   |
| 759.             | C0507E11-3 | BE824970       | ESTs<br>BE824970                                                   | C0507E11      | Mm.139860      | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 760.             | H3158D06-3 | Lnk            | linker of T-cell<br>receptor<br>pathways                           | H3158D06      | Mm.200936      | Chromosome 5             | TTGCTCT<br>GATTTTT<br>ATGGGAA<br>AAAAGTA<br>GGCAACC<br>TTAACTG<br>GTCAAAC<br>CTTGAAG<br>TCTAAGT<br>AGCCAGA<br>AACCCT<br>CAATGTG<br>TTCAATG<br>TAGTTCA<br>GACGTGC<br>ATGCCAC<br>AAATC<br>AACTGAA<br>AGTTGAC<br>AAGTGAA<br>CTTTAAC<br>ATGGAAA<br>CTTCATC<br>GGATATA<br>GTATTTC<br>AGTGATT<br>AGTGCAT<br>AAGTGCA<br>GTCTCAG<br>TAGCTTT<br>AAGAAAG<br>CTACCTA<br>GACCATT<br>AAGGAA1<br>CCCAC<br>ATTTGCA<br>CAGAAAC<br>CCAAGGT<br>CTCAGGC<br>ATCCTTA<br>GGTCTC<br>TTGGAAT<br>GGAGGAC<br>TGAAAAA<br>GTGTGTC<br>GTGTCAC<br>GCATCAT<br>TCTTATG |
| 761.             | C0174C02-3 | Pold3          | polymerase<br>(DNA-<br>directed), delta<br>3, accessory<br>subunit | C0174C02      | Mm.37562       | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 762.             | C0130G10-3 | Cklfsf7        | chemokine-like<br>factor super<br>family 7                         | C0130G10      | Mm.35600       | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 763.             | C0137F07-3 | Pik3cb         | phosphatidylo<br>sitol 3-kinase,<br>catalytic, beta<br>polypeptide | C0137F07      | Mm.213128      | Chromosome 9             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 764.             | H3115F01-3 | 2610027O18Rik  | RIKEN cDNA<br>2610027O18<br>gene                                   | H3115F01      | Mm.46501       | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 765.             | H3097F03-3 |                | Mus musculus,<br>clone<br>IMAGE:53723<br>38, mRNA                  | H3097F03      | Mm.227202      | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 766.             | H3059A05-3 | Mad2l1         | MAD2 (mitotic<br>arrest deficient,<br>homolog)-like 1<br>(yeast)   | H3059A05      | Mm.43444       | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 767.             | L0935E02-3 | Syk            | spleen tyrosine<br>kinase                                          | L0935E02      | Mm.248456      | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 768.             | C0946F08-3 | 1110014L17Rik  | RIKEN cDNA<br>1110014L17<br>gene                                   | C0946F08      | Mm.30103       | Chromosome 11            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 769.             | H3079F02-5 |                | Possibly                                                           | H3079F02      | Data not found | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                           | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A] | 60mer SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------|------------|---------------|-----------------------------------------------------|------------|------------|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            |            |               | intronic in U011488-H3079F02                        |            |            |                       | AAGTGA<br>GGATAAA<br>TAGGAAT<br>CACTCCA<br>CATGG<br>GCCTCAA<br>AACCACA<br>GGTGTGT<br>TCATCCT<br>AAAAGTC<br>TGTTTTG<br>CCGTACA<br>AGTGAAC<br>CAGCGAA<br>CCAAGGA<br>GCCATCT<br>GGCTTCT<br>AAGAAAA<br>TGTATGA<br>AGAAGAC<br>GTAATTA<br>CCCGTGT<br>GCTGTAC<br>GGCATTT<br>TTATCTT<br>TTGTAAT<br>TAAACAA<br>ACCAACC<br>TCTGTG<br>ATTAGCC<br>AGTCCGG<br>AATATTT<br>AGATCTC<br>CAGTTAG<br>AAATT<br>TACATTA<br>ATACTAA<br>ATAGAAT<br>GACTTAG<br>GTGAATA<br>ATCCTG<br>AAGATTT<br>GTCACTG<br>AAGGAAA<br>CTAAGAG<br>CGTATTG<br>CTGAGA<br>ACAAGAA<br>TTCTTAAC<br>TGAACGA<br>ATTTGCT<br>TCGATGA<br>GTTGC<br>AGGATTT<br>ATGAAGA<br>AGATGAC<br>GGTAATA<br>TAGCTGT<br>TTTCTC<br>TAGAGTC<br>AGAACAG<br>TTCAAGG<br>TTTCAATT<br>GAGTGAG<br>GAGCCA<br>TCTAAAA |
| 770.       | H3137E07-3 | Il10ra        | interleukin 10 receptor, alpha                      | H3137E07   | Mm.26658   | Chromosome 9          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 771.       | C0143H12-3 | Galns         | galactosamine (N-acetyl)-6-sulfate sulfatase        | C0143H12   | Mm.34702   | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 772.       | H3114D03-3 | Man2a1        | mannosidase 2, alpha 1                              | H3114D03   | Mm.2433    | Chromosome 17         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 773.       | H3041H09-3 | BG066348      | ESTs BG066348                                       | H3041H09   | Mm.270044  | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 774.       | C0628H04-3 | Slc2a12       | solute carrier family 2, member 12                  | C0628H04   | Mm.268014  | Chromosome 10         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 775.       | K0125E07-3 | Ifngr         | interferon gamma receptor                           | K0125E07   | Mm.549     | Chromosome 10         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 776.       | G0115E02-3 | Sdcbp         | syndecan binding protein                            | G0115E02   | Mm.276062  | Chromosome 4          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 777.       | C0032B05-3 | Rap2b         | RAP2B, member of RAS oncogene family                | C0032B05   | Mm.26939   | Chromosome 3          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 778.       | H3141C08-3 | Odf1          | oral-facial-digital syndrome 1 gene homolog (human) | H3141C08   | Mm.247480  | Chromosome X          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 779.       | H3157C05-3 | BG076236      | ESTs BG076236                                       | H3157C05   | Mm.182877  | Chromosome 1          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 780.       | H3076A01-3 | 5031439G07Rik | RIKEN cDNA                                          | H3076A01   | Mm.121973  | Chromosome 15         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                           | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------|---------------|---------------|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 5031439G07<br>gene                                                                                     |               |               |                                          | CCAAATC<br>ATGTCAC<br>AATAGGT<br>ATATACT<br>ACCCC<br>GTGTTTC<br>CATTGT<br>GTCCTGA<br>TAAATTA<br>CAGGATT<br>GACAG<br>GAAGCTG<br>GCATTGT<br>TGAAGTT<br>ATATTGA<br>TCAGCGT<br>GTCAGA<br>TTACATG<br>ATCTGAA<br>AAGACTT<br>AGGGTAA<br>AATTGAA<br>AGGAGCT<br>AGCAATC<br>TATCAAT<br>TCACACT<br>GATGAAC<br>TAAGGTA<br>ACAAGC<br>GGTGTAT<br>ATAAAGT<br>TCAATGT<br>AATCTCT<br>GTTGAAT<br>TGCTC<br>CTTTCAGT<br>CTTCTGTC<br>CGAACCT<br>CAGGATG<br>AACTTTTC<br>ACCAC |
| 781.             | H3080D06-3 | BC018507       | cDNA sequence<br>BC018507                                                                              | H3080D06      | Mm.139738     | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 782.             | L0518D04-3 | Uap1           | UDP-N-<br>acetylglucosami<br>ne<br>pyrophosphoryl<br>ase 1                                             | L0518D04      | Mm.27969      | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 783.             | K0542B11-3 | BM239901       | ESTs<br>BM239901                                                                                       | K0542B11      | Mm.222307     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 784.             | L0959D03-3 | Tnfrsfla       | tumor necrosis<br>factor receptor<br>superfamily,<br>member 1a                                         | L0959D03      | Mm.1258       | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 785.             | H3035C07-3 | BG065787       | ESTs<br>BG065787                                                                                       | H3035C07      | Mm.24933      | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 786.             | M29855.1   | Csf2rb2        | colony<br>stimulating<br>factor 2<br>receptor, beta 2,<br>low-affinity<br>(granulocyte-<br>macrophage) | M29855        | Mm.1940       | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 787.             | C0352C11-3 | BM197981       | ESTs<br>BM197981                                                                                       | C0352C11      | Mm.215584     | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 788.             | L0846B10-3 | BM117093       | ESTs<br>BM117093                                                                                       | L0846B10      | Mm.216113     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 789.             | L0227C06-3 | Serpina6a      | serine (or<br>cysteine)<br>proteinase<br>inhibitor, clade<br>B, member 6a                              | L0227C06      | Mm.2623       | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 790.             | J0214H09-3 | Serpina3g      | serine (or<br>cysteine)<br>proteinase<br>inhibitor, clade<br>A, member 3G                              | J0214H09      | Mm.264709     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------|------------|----------------|------------------------------------------------------------|---------------|---------------|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 791.             | H3077F12-3 | Arhh           | ras homolog<br>gene family,<br>member H                    | H3077F12      | Mm.20323      | Chromosome 5                             | GTTTTCA'<br>TTGGAAC<br>TTCTTTG,<br>GGCAAA(<br>GTATGAC<br>AAAATA<br>GTGTGTA<br>AATGTAA<br>GTACAAC<br>GTTTATG<br>GCTATGG<br>CAGTC<br>GTTTCCT<br>AGGTGTA<br>CGTGTC<br>GAAGCTA<br>TTATGTA'<br>AGAGA<br>TGAAAAA<br>AAAAGA/<br>GAGATGA<br>AGGAGCC<br>AGAAGTI<br>TGTCTC<br>AAAGAA/<br>AAACCGT<br>TGCGATT<br>GGGTACC<br>TAATGTA<br>GAAGTC<br>TTTCCAG<br>CTAGTTA<br>AATGAGA<br>GAAACA<br>CTATGAC<br>GGGTTTC<br>TTTTGAC<br>TTGACTG<br>AGACTGT<br>ACCTGAA<br>TCTGCTC<br>TTCCTG<br>CCCAGT<br>AACAA<br>TTTGCTA<br>TAGATCA<br>TAACAGT<br>CATT<br>GTTTTGG<br>AAAGTCG<br>GTGTCTC<br>TCCCTTC<br>GAAAAA<br>AGAGG<br>AGGAAGC<br>ATAGGCT<br>TGTATGT<br>AAGTGGA<br>ACAAGAC<br>TAGTCC<br>TACAGGG<br>GGTCTAA<br>ACCATTT<br>CACTGTA<br>TAGACAT<br>GTTGAG |
| 792.             | C0341D05-3 | BM196992       | ESTs<br>BM196992                                           | C0341D05      | Mm.222093     | Chromosome 1                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 793.             | H3043H11-3 | BG066522       | ESTs<br>BG066522                                           | H3043H11      | Mm.25035      | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 794.             | K0507D06-3 |                | Mus musculus,<br>clone<br>IMAGE:12632<br>52, mRNA          | K0507D06      | Mm.103545     | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 795.             | J0535D11-3 | AU020606       | ESTs<br>AU020606                                           | J0535D11      | Mm.26229      | Chromosome 11                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 796.             | H3152F04-3 | Sepp1          | selenoprotein P,<br>plasma, 1                              | H3152F04      | Mm.22699      | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 797.             | L0701F07-3 | H2-Ab1         | histocompatibili<br>ty 2, class II<br>antigen A, beta<br>1 | L0701F07      | Mm.275510     | Chromosome 17                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 798.             | L0227H07-3 | Clca1          | chloride<br>channel calcium<br>activated 1                 | L0227H07      | Mm.275745     | Chromosome 3                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 799.             | J1014C11-3 | 2900036G02Rik  | RIKEN cDNA<br>2900036G02<br>gene                           | J1014C11      | Mm.80676      | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 800.             | H3134H09-3 | BG074421       | ESTs<br>BG074421                                           | H3134H09      | Mm.197381     | Chromosome 12                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 801.             | G0116A07-3 | Atp6v1c1       | ATPase, H+<br>transporting,<br>V1 subunit C,<br>isoform 1  | G0116A07      | Mm.276618     | Chromosome 15                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60mer SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------|------------|---------------|--------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 802.       | L0942F05-3 | Ostm1         | osteopetrosis associated transmembrane protein 1                                                                         | L0942F05   | Mm.46636   | Chromosome 10                         | GAAACGC<br>TGTTGTA<br>TAATGAA<br>AAACTCC<br>ATTCAAT<br>AAGAA<br>AAGTTAA<br>AATACTG<br>ATCGGTC<br>AACACTC<br>AAGCTAT<br>AGCATAG<br>AAATACA<br>TTTGTAC/<br>GGCCCTG<br>TGTGAAG<br>TCTCCAT/<br>ATTAG<br>CCGTTTT/<br>ATTGGAA<br>AAGACTC<br>GAACTCA/<br>TACTGGC/<br>ATGGCA<br>GGAAAGA<br>ATCAAAC/<br>AACCTAC/<br>ATAGTTC/<br>GCCTAAG,<br>TTACTTG<br>TTGATTGC<br>TTCTGAGC<br>CAGACTC/<br>CCCTCATI<br>AATAAAAT/<br>ACATTG<br>CTAGTGA/<br>TATGTAC/<br>GACATAT/<br>ACTCTGA/<br>ATCTCTAC<br>CCACG<br>TAGTTAA1<br>TCTCTGAA<br>CATGGTA/<br>CTAGTAAC<br>GAGATACC<br>AGATTG<br>TGGATTAT<br>CGCCAAAC<br>CCCAAGTC<br>CTGTTTAA<br>GAGAAAG,<br>GAATTAA |
| 803.       | C0912H10-3 | 0610041E09Rik | RIKEN cDNA 0610041E09 gene                                                                                               | C0912H10   | Mm.132926  | Chromosome 13                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 804.       | C0304E12-3 | Pde1b         | phosphodiesterase 1B, Ca2+-calmodulin dependent                                                                          | C0304E12   | Mm.62      | Chromosome 15                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 805.       | L0605C12-3 | 4930579K19Rik | RIKEN cDNA 4930579K19 gene                                                                                               | L0605C12   | Mm.117473  | Chromosome 9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 806.       | K0539A07-3 | Cd53          | CD53 antigen                                                                                                             | K0539A07   | Mm.2692    | Chromosome 3                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 807.       | L0228H12-3 | 6430628I05Rik | RIKEN cDNA 6430628I05 gene                                                                                               | L0228H12   | Mm.196533  | Chromosome 9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 808.       | L0855B10-3 | BM117713      | ESTs BM117713                                                                                                            | L0855B10   | Mm.216997  | Chromosome 10                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 809.       | H3075B10-3 | 2810404F18Rik | RIKEN cDNA 2810404F18 gene                                                                                               | H3075B10   | Mm.29476   | Chromosome 11                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 810.       | L0022G07-3 |               | L0022G07-3 NIA Mouse E12.5 Female Mesonephros and Gonads cDNA Library Mus musculus cDNA clone L0022G07 3', MRNA sequence | L0022G07   |            | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 811.       | H3107C11-3 | Efemp2        | epidermal growth factor-containing fibulin-like extracellular matrix protein 2                                           | H3107C11   | Mm.41781   | Chromosome 19                         | GATCCAGC<br>ACCTCTGT<br>CCCTGGGC<br>ACAATGCC<br>CAGATCCG<br>TGGAAG<br>GTTCCATC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 812.       | H3025H12-3 | 1200003O06Rik | RIKEN cDNA                                                                                                               | H3025H12   | Mm.142105  | Chromosome 3                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                         | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------|------------|----------------|--------------------------------------------------------------------------------------|---------------|----------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | 1200003O06<br>gene                                                                   |               |                |                          | CTTAAAC<br>ACCGTAC<br>CAGCTCA<br>CATCCTA<br>AGAAA<br>GTAGGGC<br>AACTAAC<br>AGTAGAC<br>ATTCTAA<br>AGTAGTA<br>TGGCTTG<br>GGTGTGC<br>TATGGGC<br>CACAAAC<br>AAGAAAT<br>GGACTGC<br>TGAAAA<br>AGGTATC<br>TTACAT<br>GAATCTT<br>ACTATGT<br>AACAAAT<br>GAAGG<br>TGCTTGT<br>ACTACCT<br>ATGAAGC<br>ATGTTTA<br>TCCATAC<br>CTACTG<br>CGATGG/<br>AGATACC<br>ATGAGAC<br>GTTGAGC<br>ACAGTGC<br>TATTAT<br>GCAGCCA<br>TGGAAA/<br>AAATTAA<br>GTTGTAC<br>GTACCCA<br>AAAACC<br>TTGACAT<br>CATTACC<br>TGCAGTC<br>AATAAGC<br>ATTGTGC<br>GATAA |
| 813.             | J0040E05-3 | Stx3           | syntaxin 3                                                                           | J0040E05      | Mm.203928      | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 814.             | H3075F03-3 | C1s            | complement<br>component 1, s<br>subcomponent                                         | H3075F03      | Mm.24128       | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 815.             | L0600G09-3 | BM125147       | ESTs<br>BM125147                                                                     | L0600G09      | Mm.221784      | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 816.             | K0115H01-3 | KLHL6          | kelch-like 6                                                                         | K0115H01      | Mm.86699       | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 817.             | H3015B10-3 | Gus            | beta-<br>glucuronidase                                                               | H3015B10      | Mm.3317        | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 818.             | H3108A12-3 | 0910001A06Rik  | RIKEN cDNA<br>0910001A06<br>gene                                                     | H3108A12      | Mm.22383       | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 819.             | H3108H09-5 |                | UNKNOWN:<br>Similar to<br>Homo sapiens<br>KIAA1577<br>protein<br>(KIAA1577),<br>mRNA | H3108H09      | Data not found | Chromosome 13            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 820.             | K0645H01-3 | Fyb            | FYN binding<br>protein                                                               | K0645H01      | Mm.257567      | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 821.             | H3029A02-3 | Shyc           | selective<br>hybridizing<br>clone                                                    | H3029A02      | Mm.12912       | Chromosome 7             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 822.             | K0410D10-3 | Cxcl12         | chemokine (C-<br>X-C motif)<br>ligand 12                                             | K0410D10      | Mm.465         | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| SEQ<br>ID<br>NO: | CLONE ID    | GENE<br>SYMBOL | GENE<br>NAME                                               | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------|-------------|----------------|------------------------------------------------------------|---------------|---------------|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 823.             | H3118H11-3  | Snrpg          | small nuclear<br>ribonucleoprote<br>in polypeptide<br>G    | H3118H11      | Mm.21764      | Chromosome 18                   | CATGAGC<br>GCCCCAC<br>CGAGCTC<br>AAGTTT/<br>CAAGAA/<br>CATTGA/<br>CTCTGTA<br>CAAGTT/<br>GCATTT/<br>TAATTAT<br>AAGTCC/<br>CTGGC<br>TTTTCAC<br>ATAAAA/<br>ATGTGG/<br>AGGCAT/<br>CCACCG/<br>TACCAC<br>AAGAAG<br>GGAAAA<br>GAGAGT/<br>AACCGC/<br>GAACTA/<br>TCTGCTC<br>CCTGATC<br>CTGTGT/<br>AGGAGG<br>AGTTAT/<br>ATTCTCA<br>AGGAAA<br>AGCAAA/<br>CATTTTA<br>AAGTAC/<br>TTATTTT<br>GTCCTGC<br>GGGGGT<br>CTGCAC/<br>TGGACTC<br>ACTTGC/<br>TATCTAC<br>ACAAGA/<br>ATGCTAC<br>AGATTTC<br>TACTTTC<br>GGTGT/<br>AAGGCC/<br>GTTGCA/<br>TTGCAC<br>ATAAAA/<br>AAACTA/<br>ATGCTT/<br>TGCACA/<br>AGTATA/<br>GATGGG<br>ACCTAA/<br>CATGAC/<br>ACTATT/<br>GCTATA/<br>TGAACC/<br>TGTGC<br>TTTATAC<br>AGGTTT/<br>AGAGAG<br>TAATTT/<br>CAGCCT/<br>TGTTC |
| 824.             | K0517D08-3  | BM238427       | ESTs<br>BM238427                                           | K0517D08      | Mm.222266     | Chromosome 19                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 825.             | L0227G11-3  | Sh3d1B         | SH3 domain<br>protein 1B                                   | L0227G11      | Mm.40285      | Chromosome 12                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 826.             | H3134B10-3  | 6530409L22Rik  | RIKEN cDNA<br>6530409L22<br>gene                           | H3134B10      | Mm.41940      | Chromosome Multiple<br>Mappings |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 827.             | H3115A08-3  | Ly6a           | lymphocyte<br>antigen 6<br>complex, locus<br>A             | H3115A08      | Mm.263124     | Chromosome 15                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 828.             | C0120G03-3  | Csk            | c-src tyrosine<br>kinase                                   | C0120G03      | Mm.21974      | Chromosome 9                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 829.             | H3094G08-3  | Tigd2          | tigger<br>transposable<br>element derived<br>2             | H3094G08      | Mm.25843      | Chromosome 6                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 830.             | NM_008362.1 | Il1r1          | interleukin 1<br>receptor, type I                          | NM_008362     | Mm.896        | Chromosome 1                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 831.             | C0300E10-3  | Trps1          | trichorhinophal<br>angeal<br>syndrome 1<br>(human)         | C0300E10      | Mm.30466      | Chromosome 15                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 832.             | L0274A03-3  | Ptpn2          | protein tyrosine<br>phosphatase,<br>non-receptor<br>type 2 | L0274A03      | Mm.260433     | Chromosome 18                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 833.             | H3005H07-3  | 1810031K02Rik  | RIKEN cDNA<br>1810031K02<br>gene                           | H3005H07      | Mm.145384     | Chromosome 4                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                                       | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]        | 60mer<br>SEQUENCE                                                                                                                                                                                                                                               |
|------------------|------------|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 834.             | H3109H12-3 | 1810009M01Rik  | RIKEN cDNA<br>1810009M01<br>gene                                                                                                                                   | H3109H12      | Mm.28385      | Chromosome Multiple<br>Mappings | TTCTTCC/<br>ACAGATA<br>GTCATTT<br>CAATGCC<br>AAGGAG/<br>AACTTG<br>TACGTGG<br>GGGACCT<br>TTGGAAT<br>TTGTTGT<br>AAAAC TG<br>AAAGGA<br>ACCAACT<br>TCAAAGA<br>GTAAAGA<br>GAGATAC<br>ATCTTTG<br>ATAGTC                                                               |
| 835.             | J0008D01-3 | Enpp1          | ectonucleotide<br>pyrophosphatas<br>e/phosphodiesterase 1                                                                                                          | J0008D01      | Mm.27254      | Chromosome 10                   | TTCTTCC/<br>ACAGATA<br>GTCATTT<br>CAATGCC<br>AAGGAG/<br>AACTTG<br>TACGTGG<br>GGGACCT<br>TTGGAAT<br>TTGTTGT<br>AAAAC TG<br>AAAGGA<br>ACCAACT<br>TCAAAGA<br>GTAAAGA<br>GAGATAC<br>ATCTTTG<br>ATAGTC                                                               |
| 836.             | H3119H05-3 | Mafb           | v-maf<br>musculoaponeurotic<br>fibrosarcoma<br>oncogene<br>family, protein<br>B (avian)                                                                            | H3119H05      | Mm.233891     | Chromosome 10                   | TTCTTCC/<br>ACAGATA<br>GTCATTT<br>CAATGCC<br>AAGGAG/<br>AACTTG<br>TACGTGG<br>GGGACCT<br>TTGGAAT<br>TTGTTGT<br>AAAAC TG<br>AAAGGA<br>ACCAACT<br>TCAAAGA<br>GTAAAGA<br>GAGATAC<br>ATCTTTG<br>ATAGTC                                                               |
| 837.             | H3048G11-3 | Blvrb          | biliverdin<br>reductase B<br>(flavin<br>reductase<br>(NADPH))                                                                                                      | H3048G11      | Mm.24021      | Chromosome 7                    | TGACACA<br>GAGGGGT<br>TAAATTT<br>CCAAAAG<br>AAATTCT<br>GGAAGC<br>ATCACCA<br>TAGTGTC<br>TCATTGT<br>AACGCTC<br>ACCTTCA<br>TAATAG<br>GCCGCTT<br>GTAACCT<br>GGCCCCA<br>TAAGGGC<br>GTTTTGG<br>TTGTA<br>CCAAGAA<br>GTATAAA<br>AGCTCTG<br>ACTGAAA<br>TTCAAGT<br>TCGATC |
| 838.             | H3107D05-3 | 1110004C05Rik  | RIKEN cDNA<br>1110004C05<br>gene                                                                                                                                   | H3107D05      | Mm.141021     | Chromosome 7                    | TGACACA<br>GAGGGGT<br>TAAATTT<br>CCAAAAG<br>AAATTCT<br>GGAAGC<br>ATCACCA<br>TAGTGTC<br>TCATTGT<br>AACGCTC<br>ACCTTCA<br>TAATAG<br>GCCGCTT<br>GTAACCT<br>GGCCCCA<br>TAAGGGC<br>GTTTTGG<br>TTGTA<br>CCAAGAA<br>GTATAAA<br>AGCTCTG<br>ACTGAAA<br>TTCAAGT<br>TCGATC |
| 839.             | H3006B01-3 | Cklfs3         | chemokine-like<br>factor super<br>family 3                                                                                                                         | H3006B01      | Mm.292081     | Chromosome 8                    | AGGACAT<br>CAACTTC<br>CAATAAT<br>GATTTCC<br>GACAAAT<br>ACAAGTG<br>GGGGAGT<br>ATAATAG<br>ATTCATA<br>CAAGAAC<br>AAAATGG<br>GACTTT<br>TGCCACT<br>CTGACTTC<br>AATATGG<br>TTAAACA<br>AAAGTGA<br>TTTAA                                                                |
| 840.             | L0853H04-3 |                | transcribed<br>sequence with<br>weak similarity<br>to protein<br>pir:A43932<br>(H.sapiens)<br>A43932 mucin<br>2 precursor,<br>intestinal -<br>human<br>(fragments) | L0853H04      | Mm.275315     | Chromosome 12                   | AGGACAT<br>CAACTTC<br>CAATAAT<br>GATTTCC<br>GACAAAT<br>ACAAGTG<br>GGGGAGT<br>ATAATAG<br>ATTCATA<br>CAAGAAC<br>AAAATGG<br>GACTTT<br>TGCCACT<br>CTGACTTC<br>AATATGG<br>TTAAACA<br>AAAGTGA<br>TTTAA                                                                |
| 841.             | C0949G05-3 | BM221093       | ESTs<br>BM221093                                                                                                                                                   | C0949G05      | Mm.221696     | Chromosome 6                    | AGGACAT<br>CAACTTC<br>CAATAAT<br>GATTTCC<br>GACAAAT<br>ACAAGTG<br>GGGGAGT<br>ATAATAG<br>ATTCATA<br>CAAGAAC<br>AAAATGG<br>GACTTT<br>TGCCACT<br>CTGACTTC<br>AATATGG<br>TTAAACA<br>AAAGTGA<br>TTTAA                                                                |
| 842.             | K0648D10-3 | Tlr1           | toll-like<br>receptor 1                                                                                                                                            | K0648D10      | Mm.33922      | Chromosome 5                    | AGGACAT<br>CAACTTC<br>CAATAAT<br>GATTTCC<br>GACAAAT<br>ACAAGTG<br>GGGGAGT<br>ATAATAG<br>ATTCATA<br>CAAGAAC<br>AAAATGG<br>GACTTT<br>TGCCACT<br>CTGACTTC<br>AATATGG<br>TTAAACA<br>AAAGTGA<br>TTTAA                                                                |
| 843.             | H3014E09-3 | BC017643       | cDNA sequence<br>BC017643                                                                                                                                          | H3014E09      | Mm.27182      | Chromosome 11                   | AGGACAT<br>CAACTTC<br>CAATAAT<br>GATTTCC<br>GACAAAT<br>ACAAGTG<br>GGGGAGT<br>ATAATAG<br>ATTCATA<br>CAAGAAC<br>AAAATGG<br>GACTTT<br>TGCCACT<br>CTGACTTC<br>AATATGG<br>TTAAACA<br>AAAGTGA<br>TTTAA                                                                |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                                              | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]        | 60m SEQUE                                                                                                                                                                                                                                                      |
|------------|------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 844.       | H3022D06-3 | Il2rg         | interleukin 2 receptor, gamma chain                                                                                                                    | H3022D06   | Mm.2923    | Chromosome X                 | CATCAAT<br>TGATGGA<br>CAAAGTC<br>AGTCCTA<br>ACGCTAA<br>CCCTA<br>CAGTTGG<br>AATGGAT<br>GCTCAAT<br>AAGAGGC<br>ATACAGC<br>ACTCTGG<br>TCAGTCA<br>TGCATAA<br>AAATCAA<br>AAGAGCT<br>AAGGTTA<br>AGGTCA                                                                |
| 845.       | L0201A03-3 | 2410004H05Rik | RIKEN cDNA 2410004H05 gene                                                                                                                             | L0201A03   | Mm.8766    | Chromosome 14                |                                                                                                                                                                                                                                                                |
| 846.       | H3026E03-5 |               | Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430039 C10 product:unkno wn EST, full insert sequence | H3026E03   | Mm.249306  | Chromosome Un: not placed    | abhydrolase domain containing 2                                                                                                                                                                                                                                |
| 847.       | H3091E12-3 | Abhd2         |                                                                                                                                                        | H3091E12   | Mm.87337   | Chromosome 7                 | AGCAGGT<br>CGGACTT<br>TGAGCAA<br>ATTTTTTC<br>ATATGAG<br>TTTAC<br>CTTGCTTC<br>AGCAAAA<br>CTGGTTTC<br>AAGAGGA<br>CTGTCCA<br>GGCCCC<br>TCTCAAT<br>AAGGTGT<br>CCTATCA<br>ACTTGAA<br>ATATGGT<br>ACCCA                                                               |
| 848.       | H3003E01-3 | Cutl1         | cut-like 1 (Drosophila)                                                                                                                                | H3003E01   | Mm.258225  | Chromosome 5                 |                                                                                                                                                                                                                                                                |
| 849.       | H3016H08-5 | Crsp9         | cofactor required for Sp1 transcriptional activation, subunit 9, 33kDa                                                                                 | H3016H08   | Mm.24159   | Chromosome 11                |                                                                                                                                                                                                                                                                |
| 850.       | C0118E09-3 | Oas1a         | 2'-5' oligoadenylate synthetase 1A                                                                                                                     | C0118E09   | Mm.14301   | Chromosome 5                 | ACTGGAC<br>GTATTATC<br>TTCAACAC<br>GAGGTCTC<br>ATACCTGC<br>GACAGC<br>GGCTGTT<br>GTAAAATC<br>TTTGTGTT<br>TTACAAC<br>GCTTTTAC<br>CACAG<br>TGAGTGC<br>TGTCAGA<br>ACCAAGA<br>CTCCAAGC<br>GTAGGTA<br>GTGGTT<br>GTCATTGT<br>AGGTGAC<br>AGGAACT<br>CGTTAAA<br>CGAGCCT |
| 851.       | L0535B02-3 | Col15a1       | procollagen, type XV                                                                                                                                   | L0535B02   | Mm.233547  | Chromosome 4                 |                                                                                                                                                                                                                                                                |
| 852.       | L0500E02-3 | Sgcg          | sarcoglycan, gamma (dystrophin-associated glycoprotein)                                                                                                | L0500E02   | Mm.72173   | Chromosome 14                |                                                                                                                                                                                                                                                                |
| 853.       | H3077B08-3 | 5330431K02Rik | RIKEN cDNA 5330431K02 gene                                                                                                                             | H3077B08   | Mm.101992  | Chromosome Multiple Mappings |                                                                                                                                                                                                                                                                |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                          | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60mer<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------|------------|----------------|-------------------------------------------------------|---------------|---------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 854.             | J0209G02-3 | Gnb4           | guanine<br>nucleotide<br>binding protein,<br>beta 4   | J0209G02      | Mm.9336       | Chromosome 3             | TCATGA<br>TCTTAGA<br>GGAATTG<br>CCATATT<br>GTTCTCC<br>ATACCTG<br>AATCC<br>TGCTTTCT<br>TCTTTAA<br>ATTTATT<br>TCTCATT<br>TAAAACC<br>GTATT<br>CTGCATG<br>AACTTTA<br>ATGGTGT<br>CATATAA<br>TGAGAAT<br>TATAC<br>CGTGCTG<br>ACGAGAG<br>CAGAAGC<br>GAAGCAA<br>GAGAAGC<br>CTGAACA<br>TGGAGGC<br>TACCCAA<br>TTTCAAG<br>AAGGAAA<br>AGAAGT<br>GATTACA<br>TGGAGGA<br>TGTGAAA<br>AAGTCAC<br>ACAAACC<br>GTGCCTA<br>CTCTGTC<br>CTATTTTG<br>AGACATC<br>TGCCTGA<br>ACTGTGG<br>ATCCAAA<br>GTCCA<br>TAATTATC<br>ATTGGGG<br>TGAAGTAC<br>AGATCCA<br>AACTACG<br>TCTCCG<br>TTGGGTAT<br>TTATGTTT<br>TCATAACA<br>GCAATAAC<br>TAGGAAA<br>TACCG<br>TCTGATGT<br>AGTGCGG<br>TCCTGGTT<br>CTCACAGC<br>TTTTAATT<br>CTAAG<br>ATCTCCTG<br>ATGTATTT<br>TCAAATGC<br>GCCTTAAT<br>GAAATCTG |
| 855.             | C0661E01-3 | Lcn7           | lipocalin 7                                           | C0661E01      | Mm.15801      | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 856.             | K0221E09-3 | Scml2          | sex comb on<br>midleg-like 2<br>(Drosophila)          | K0221E09      | Mm.159173     | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 857.             | C0184F12-3 | D8Ert594e      | DNA segment,<br>Chr 8, ERATO<br>Doi 594,<br>expressed | C0184F12      | Mm.235074     | Chromosome 8             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 858.             | L0602B03-3 | Myoz2          | myozenin 2                                            | L0602B03      | Mm.141157     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 859.             | C0944F04-3 | 1110055E19Rik  | RIKEN cDNA<br>1110055E19<br>gene                      | C0944F04      | Mm.39046      | Chromosome 6             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 860.             | L0004A03-3 | Gli2           | GLI-Kruppel<br>family member<br>GLI2                  | L0004A03      | Mm.12090      | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 861.             | L0860B03-3 |                | ESTs<br>AV321020                                      | L0860B03      | Mm.221891     | Chromosome 5: not placed |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 862.             | L0841F10-3 | 2310045A20Rik  | RIKEN cDNA<br>2310045A20<br>gene                      | L0841F10      | Mm.235020     | Chromosome 5             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 863.             | L0008H10-3 | Agri           | agrin                                                 | L0008H10      | Mm.269006     | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 864.             | C0128B02-3 | Casq1          | calsequestrin 1                                       | C0128B02      | Mm.12829      | Chromosome 1             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                                                      | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A]                 | 60m SEQUE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------|------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 865.       | C0645C09-3 | BM209340      | ESTs<br>BM209340                                                                                                                               | C0645C09   | Mm.222131  | No Chromosome location info available | GCAGAA<br>GCAGCA/<br>AAAAGA/<br>GAGAGC/<br>GGCAAG/<br>CTCTCTG<br>TCCCTTT<br>TGAGGAA/<br>CCCCATG<br>ACCTTAT<br>CTAAGAC<br>CGTGATC<br>AGTCGT<br>ACCGGCT<br>CCAAATA<br>CGTCATT<br>TATGAAG<br>TCAGCCC<br>AGATTT                                                                                                                                                                                                                                                                                          |
| 866.       | H3082B03-3 | Mylk          | myosin, light polypeptide kinase                                                                                                               | H3082B03   | Mm.288200  | Chromosome 16                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 867.       | C0309D09-3 |               | transcribed sequence with moderate similarity to protein sp:P00722 (E. coli)<br>BGAL_ECOLI<br>Beta-galactosidase (Lactase)<br>ESTs<br>BG076287 | C0309D09   | Mm.213420  | Chromosome 11                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 868.       | H3157H09-3 | BG076287      | ESTs<br>BG076287                                                                                                                               | H3157H09   | Mm.131026  | Chromosome 2                          | ATGGTTT/<br>CAGCAAT<br>CATTGCC/<br>GGGTCTA<br>GAATAAG<br>TTCTTG<br>ACAATCT/<br>CAGCGAA<br>TTCTACA/<br>CTGTGCT/<br>AACATGT/<br>TCCAAG<br>AACTGTT/<br>GATTGAA.<br>CCATCCCC<br>CCCTAAA/<br>GTGCCTT/<br>AACCC<br>CGACTGA/<br>ATGACAT/<br>AGACTTT/<br>TATGCTGC<br>GAATGAA/<br>GAGATA<br>TGCCCTCT<br>CGCCAGT/<br>CAAAGGG/<br>AGAGAGC/<br>CTAGCAG/<br>TAGTGTT<br>CCACTAA/<br>TAGCCAG/<br>CATGTAG/<br>ACACATG/<br>ACACAGA/<br>AAACTTTT<br>AGAAATG/<br>ATACATTG<br>GCATTAG<br>TAAGTTGT<br>GACAGGG/<br>TTAAGTG |
| 869.       | H3061D03-3 | Pcsk5         | proprotein convertase subtilisin/kexin type 5                                                                                                  | H3061D03   | Mm.3401    | Chromosome 19                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 870.       | L0843D01-3 | 3732412D22Rik | RIKEN cDNA 3732412D22 gene                                                                                                                     | L0843D01   | Mm.18830   | No Chromosome location info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 871.       | L0702H07-3 | 5830415L20Rik | RIKEN cDNA 5830415L20 gene                                                                                                                     | L0702H07   | Mm.46184   | Chromosome 5                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 872.       | L0548G08-3 | Xin           | cardiac morphogenesis                                                                                                                          | L0548G08   | Mm.10117   | Chromosome 9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 873.       | L0803E02-3 | Nkd1          | naked cuticle 1 homolog (Drosophila)                                                                                                           | L0803E02   | Mm.30219   | Chromosome 8                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 874.       | C0925G12-3 | Fbxo30        | F-box protein 30                                                                                                                               | C0925G12   | Mm.276229  | Chromosome 10                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

| SEQ ID NO: | CLONE ID   | GENE SYMBOL   | GENE NAME                                                                                                              | CLONE NAME | UG CLUSTER | CHR_LOCATION PENG [A] | 60m SEQUE                                                                                                                                                                                                                                                                                                                                                                   |
|------------|------------|---------------|------------------------------------------------------------------------------------------------------------------------|------------|------------|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 875.       | L0911A11-3 | 2010313D22Rik | RIKEN cDNA 2010313D22 gene                                                                                             | L0911A11   | Mm.260594  | Chromosome 5          | CAAACG<br>CCTGTCT<br>CTTTTCT<br>GAATTT<br>AGGAAA<br>TGTAGC<br>ACCGTTC<br>ACTGTGC<br>AGAAGA<br>TCACTA<br>CTATGAC<br>GGGAAG<br>CTATTTT<br>AGATGT<br>GCGGAG<br>GTAATA<br>CCAGAG<br>CTATAG<br>ACCCAA<br>GTGCTCT<br>CTTTTAC<br>GGATTT<br>CATGTGC<br>AAAAAT<br>TTACCA<br>GGTTAA<br>CAAATT<br>AATAAC<br>TTGAAT<br>GCAGG<br>TCACCA<br>TGAAAG<br>ACTACC<br>TTAACA<br>GATTTA<br>CTCAG |
| 876.       | AF084466.1 | Rrad          | Ras-related associated with diabetes                                                                                   | AF084466   | Mm.29467   | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                             |
| 877.       | H3073G09-3 | 1600029N02Rik | RIKEN cDNA 1600029N02 gene                                                                                             | H3073G09   | Mm.154121  | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                                             |
| 878.       | L0815B08-3 | 1100001D19Rik | RIKEN cDNA 1100001D19 gene                                                                                             | L0815B08   | Mm.260515  | Chromosome X          |                                                                                                                                                                                                                                                                                                                                                                             |
| 879.       | J1037H05-3 | D230016N13Rik | RIKEN cDNA D230016N13 gene                                                                                             | J1037H05   | Mm.21686   | Chromosome 13         |                                                                                                                                                                                                                                                                                                                                                                             |
| 880.       | K0421F09-3 |               | transcribed sequence with weak similarity to protein ref:NP_081764.1 (M.musculus) RIKEN cDNA 5730493B19 [Mus musculus] | K0421F09   | Mm.222196  | Chromosome 6          |                                                                                                                                                                                                                                                                                                                                                                             |
| 881.       | H3082E06-3 | 1110003B01Rik | RIKEN cDNA 1110003B01 gene                                                                                             | H3082E06   | Mm.275648  | Chromosome 13         | TGTTGCC<br>GATATG<br>TCAACT<br>GGAAAG<br>CTACTC<br>AGGAC<br>TCTAAC<br>TATTTG<br>TCTTTA<br>GAACAA<br>TCTTGA<br>TAAAT<br>CGACAC<br>GGCCCT<br>AGGTAG<br>CATCTA<br>ATCTGG<br>AAGCTC<br>TCTCAG<br>TTGAAG<br>TCATCT<br>CCTCCA<br>ACAGAT<br>CCCAA<br>TCTTTTC<br>CGATCA<br>ATGAGT                                                                                                   |
| 882.       | C0935B04-3 | Hhip          | Hedgehog-interacting protein                                                                                           | C0935B04   | Mm.254493  | Chromosome 8          |                                                                                                                                                                                                                                                                                                                                                                             |
| 883.       | H3116B02-3 | 1110007C05Rik | RIKEN cDNA 1110007C05 gene                                                                                             | H3116B02   | Mm.27571   | Chromosome 7          |                                                                                                                                                                                                                                                                                                                                                                             |
| 884.       | C0945G10-3 | Tp53i11       | tumor protein p53 inducible protein 11                                                                                 | C0945G10   | Mm.41033   | Chromosome 2          |                                                                                                                                                                                                                                                                                                                                                                             |
| 885.       | K0440G09-3 | Tgfb3         | transforming growth factor, beta 3                                                                                     | K0440G09   | Mm.3992    | Chromosome 12         |                                                                                                                                                                                                                                                                                                                                                                             |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                   | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                              |
|------------------|------------|----------------|----------------------------------------------------------------|---------------|---------------|------------------------------------------|----------------------------------------------------------------------------------------------|
| 886.             | L0916G12-3 | BM118833       | ESTs<br>BM118833                                               | L0916G12      | Mm.221415     | Chromosome 6                             | CAGATCA/<br>ATTAGTT<br>GGCCA<br>TGGGAA/<br>TTTAGGA<br>ATTGTAT<br>TTTGCAA<br>CATAAGC<br>ATGCC |
| 887.             | L0505A04-3 | Dnajb5         | DnaJ (Hsp40)<br>homolog,<br>subfamily B,<br>member 5           | L0505A04      | Mm.20437      | Chromosome 4                             | TACTCCC<br>TTGTATA<br>TCGAAT/<br>AGGAGC/<br>AGAAAA/<br>TCAGCT                                |
| 888.             | L0542E08-3 | Usmg4          | upregulated<br>during skeletal<br>muscle growth<br>4           | L0542E08      | Mm.27881      | Chromosome 3                             | CCGCAC1<br>CTAGCAC<br>CTTACAT<br>TCAAAGT1<br>CGACTTC<br>ACTCT                                |
| 889.             | L0223E12-3 | Sparcl1        | SPARC-like 1<br>(mast9, hevin)                                 | L0223E12      | Mm.29027      | Chromosome 5                             | GCTTTGG<br>AAAGAG/<br>ATATAGA/<br>AACCCCC<br>TTGAATT                                         |
| 890.             | K0349C07-3 | 4631423F02Rik  | RIKEN cDNA<br>4631423F02<br>gene                               | K0349C07      | Mm.68617      | Chromosome 1                             | TTTGAG<br>AAATCAC<br>GCAGGTG<br>GATAAA/<br>TAATGTT<br>ATTCGGC                                |
| 891.             | C0302A11-3 |                | EST BI988881                                                   | C0302A11      | Mm.260361     | No Chromosome location<br>info available | CTCAC<br>GAACCA/<br>TGGAATC<br>CATAAG/<br>TCAACAC<br>CCTCTCA                                 |
| 892.             | C0930C11-3 | Fgf13          | fibroblast<br>growth factor<br>13                              | C0930C11      | Mm.7995       | Chromosome X                             | TGTATG<br>GTATCGT<br>CCCAGTC<br>AGATAAC<br>AACAAAG/<br>CCTCAAC                               |
| 893.             | H3022A11-3 | Cald1          | caldesmon 1                                                    | H3022A11      | Mm.130433     | Chromosome 6                             | GATTT<br>GTCAAA/<br>CCTTCAG<br>CCTTAGA<br>CAGAAAG/<br>TTGATCC                                |
| 894.             | C0660B06-3 | Csrp1          | cysteine and<br>glycine-rich<br>protein 1                      | C0660B06      | Mm.196484     | Chromosome 1                             | ATAACAC<br>AATAGA/<br>TTCACCT<br>ATGGAG/<br>AGCCAG/<br>AGGACCC                               |
| 895.             | L0949F12-3 | Hey1           | hairy/enhancer-<br>of-split related<br>with YRPW<br>motif-like | L0949F12      | Mm.103615     | Chromosome 4                             | AGTCTAC<br>CGTGGAC<br>GGGCTAC<br>AGCTCTC<br>TAATCTT<br>ACATAC1                               |
| 896.             | K0225B06-3 | Unc5c          | unc-5 homolog<br>C (C. elegans)                                | K0225B06      | Mm.24430      | Chromosome 3                             | AATGAG<br>CTTATAG<br>AATGTTC<br>CCTCAAT                                                      |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                              | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A]                 | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------|------------|----------------|-------------------------------------------|---------------|---------------|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 897.             | K0541E04-3 | Herc3          | hect domain<br>and RLD 3                  | K0541E04      | Mm.33788      | Chromosome 6                             | ACTCAT<br>CAGTATC<br>CTGGA<br>AGCAGG<br>TTATGTT<br>CAAATG<br>GTCTCA<br>GACATG<br>CTGCTC<br>ACTCTG<br>TACTGG<br>CTCTGT<br>GACAAA<br>ATGTGCC<br>CAGTA<br>TTACAG<br>TGTTGT<br>TTTGTG<br>GCTTCC<br>AGAATT<br>GATAC<br>ATGGAA<br>GGTCAT<br>GAACAT<br>GATCTT<br>ACAAGT<br>TGTTAA<br>TAAAT<br>TCCTGG<br>ATGACC<br>AACTTC<br>TTATTTC<br>GGGAA<br>TCGACG<br>CTTACCT<br>AGGCAA<br>TATCCC<br>GATCAG<br>CCCAA<br>ACCTGT<br>GTTTTTC<br>AAGAAA<br>AGTGCA<br>GATAGC<br>CTTGAG<br>CTGCAG<br>TCTCAT<br>GAAAAA<br>CTACAA<br>AAACAG<br>CATGGG<br>AAAGAT<br>CCACGT<br>GTAGTG<br>ACCCGA<br>AATATG<br>ATCTTT<br>GTGTTG<br>TAATTT<br>TAAAGT<br>AGTAGG<br>GTAAAT<br>GACTAT<br>CTTGGG<br>GCACTC<br>ACATGG |
| 898.             | C0151A03-3 | BC026744       | cDNA sequence<br>BC026744                 | C0151A03      | Mm.4079       | Chromosome 5                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 899.             | L0045C07-3 | 6-Sep          | septin 6                                  | L0045C07      | Mm.258618     | Chromosome X                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 900.             | L0509E03-3 | Ryr2           | ryanodine<br>receptor 2,<br>cardiac       | L0509E03      | Mm.195900     | Chromosome 13                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 901.             | H3049B08-3 | Tes            | testis derived<br>transcript              | H3049B08      | Mm.271829     | Chromosome 6                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 902.             | L0533C09-3 | BM123974       | ESTs<br>BM123974                          | L0533C09      | Mm.213265     | Chromosome 14                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 903.             | H3108C01-3 | 4930444A02Rik  | RIKEN cDNA<br>4930444A02<br>gene          | H3108C01      | Mm.17631      | Chromosome 8                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 904.             | C0110C06-3 | Epb4.111       | erythrocyte<br>protein band<br>4.1-like 1 | C0110C06      | Mm.20852      | Chromosome 2                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 905.             | C0324H08-3 | Enah           | enabled<br>homolog<br>(Drosophila)        | C0324H08      | Mm.87759      | Chromosome Multiple<br>Mappings          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 906.             | C0917A09-3 |                | ESTs<br>BB231855                          | C0917A09      | Mm.242207     | No Chromosome location<br>info available |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 907.             | L0854B10-3 | Anks1          | ankyrin repeat<br>and SAM<br>domain       | L0854B10      | Mm.32556      | Chromosome 17                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                   | CLONE<br>NAME | UG<br>CLUSTER | CHR_LOCATION<br>PENG [A] | 60m<br>SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------|------------|----------------|----------------------------------------------------------------|---------------|---------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |            |                | containing 1                                                   |               |               |                          | ATCATCA/<br>AGTTCAC<br>AGCTT<br>CCCTAAC<br>TGAAAC/<br>ACTCTGT<br>CCTGTGC<br>ATTTAA/<br>AAATG<br>ATTTATA<br>TATCCTT<br>TGCTGAC<br>GTAACTC<br>TGTTTCT<br>AAGTC<br>ACCTGT/<br>CACTGTC<br>TGTGGGC<br>CTGGTC/<br>AACTTGT<br>ATAAA<br>TAATCCC<br>AAAGTC/<br>CTGTGGC<br>TAGAAC/<br>ACTCACT<br>CTGGTA<br>TTAGCTT<br>ACCCCA/<br>AAGGTT/<br>AACAAG/<br>TGCCTG/<br>TACTT<br>AATAAA/<br>CCTTAG/<br>ACTGTA/<br>CTTCAA/<br>TCATGTA<br>TAGGCA<br>AGAGAT/<br>ACTACAC<br>TAGATTC<br>TTTTAGT<br>ATTAATC<br>GGAGTA<br>TATGGCC<br>GGTTTCA<br>GTCAGG/<br>TCTAATC<br>GTGGCA/<br>AGCAA<br>TGTGTCA<br>AATCCTC<br>AACCTGC<br>TTAATCC<br>GGACCTC<br>TGGAG<br>CCACCCA/<br>AATGAC/<br>AAGTAG/<br>CAGTTTA<br>AGTTAG/<br>TTCTAC<br>CATAGTC<br>ATATGCT<br>TTTTATG |
| 908.             | K0326D08-3 | Ly75           | lymphocyte<br>antigen 75                                       | K0326D08      | Mm.2074       | Chromosome 2             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 909.             | H3074H01-3 | C430017H16     | hypothetical<br>protein<br>C430017H16                          | H3074H01      | Mm.268854     | Chromosome 3             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 910.             | H3131D02-3 | Tnk2           | tyrosine kinase,<br>non-receptor, 2                            | H3131D02      | Mm.1483       | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 911.             | C0112B03-3 | Heyl           | hairy/enhancer-<br>of-split related<br>with YRPW<br>motif-like | C0112B03      | Mm.103615     | Chromosome 4             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 912.             | L0514A09-3 | 6430511F03     | hypothetical<br>protein<br>6430511F03                          | L0514A09      | Mm.19738      | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 913.             | C0234D07-3 | Fbxo30         | F-box protein<br>30                                            | C0234D07      | Mm.276229     | Chromosome 10            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 914.             | H3152A02-3 | St6gal1        | beta galactoside<br>alpha 2,6<br>sialyltransferase<br>1        | H3152A02      | Mm.149029     | Chromosome 16            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 915.             | H3075C04-3 | Ches1          | checkpoint<br>suppressor 1                                     | H3075C04      | Mm.268534     | Chromosome 12            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 916.             | L0600E02-3 | BM125123       | ESTs<br>BM125123                                               | L0600E02      | Mm.221782     | Chromosome 19            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 917.             | K0501F10-3 | BM237456       | ESTs<br>BM237456                                               | K0501F10      | Mm.34527      | Chromosome X             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 918.             | K0301H08-3 | Oxct           | 3-oxoacid CoA<br>transferase                                   | K0301H08      | Mm.13445      | Chromosome 15            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| SEQ<br>ID<br>NO: | CLONE ID   | GENE<br>SYMBOL | GENE<br>NAME                                                                                                                                   | CLONE<br>NAME | UG<br>CLUSTER  | CHR_LOCATION<br>PENG [A]              | 60m<br>SEQUENCE                                                                                                                                                                 |
|------------------|------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------|---------------|----------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 919.             | L0229E07-3 | Lu             | Lutheran blood group (Aubergier b antigen included)                                                                                            | L0229E07      | Mm.29236       | Chromosome 7                          | ATGTATT<br>CTCGACT<br>CCTGAA<br>GTTGAGC<br>CGACCTC<br>AGGCAA<br>GGATCTC<br>TTTGGGC<br>TCGGA                                                                                     |
| 920.             | H3077C06-3 | 4931430I01Rik  | RIKEN cDNA 4931430I01 gene                                                                                                                     | H3077C06      | Mm.12454       | Chromosome 1                          | ACCAAC<br>GACTAG<br>TGCTATC<br>CCTGTCT<br>GCTCTTA<br>TGCCTA                                                                                                                     |
| 921.             | J0807D02-3 |                | Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930022I23 product:unclassified, full insert sequence. | J0807D02      | Mm.125975      | Chromosome 7                          | CCAGGG<br>AACGAT<br>CAGTGG<br>AAATAT<br>CCTCAAC<br>AAAGAT                                                                                                                       |
| 922.             | H3118G11-3 | C130068N17     | hypothetical protein C130068N17                                                                                                                | H3118G11      | Mm.138073      | Chromosome 2                          | GGTGCA<br>GTACTC<br>GTCACA<br>ACGCAT<br>AAGGTA<br>CTAAAT<br>AGATCA<br>CTGGAC<br>GATCCA<br>CGATTG<br>ATAAAC<br>CAAGA                                                             |
| 923.             | L0818F01-3 | Smardc3        | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3                                              | L0818F01      | Mm.140672      | Chromosome 5                          | ATATCC<br>AACTTA<br>AGTTAG<br>TTGTTA<br>AAAAAT<br>GTCTGG<br>AAAGCA<br>TTAGTA<br>CTGGTG<br>AGTCTT<br>ATTGAT<br>TTTTCC<br>CAACTT<br>ATAATG<br>CATTGA<br>CATTTG<br>GGTTAT<br>GGGAT |
| 924.             | C0359A10-3 | BM198389       | ESTs BM198389                                                                                                                                  | C0359A10      | Mm.218312      | Chromosome 1                          | AGGAAT<br>ACGTTT<br>AAGTAA<br>TTACAG<br>AAGTGT<br>GCTCA                                                                                                                         |
| 925.             | G0108E12-3 | 1190009E20Rik  | RIKEN cDNA 1190009E20 gene                                                                                                                     | G0108E12      | Mm.260102      | Chromosome 5                          |                                                                                                                                                                                 |
| 926.             | C0941C09-3 | Gja7           | gap junction membrane channel protein alpha 7                                                                                                  | C0941C09      | Mm.3096        | Chromosome 11                         |                                                                                                                                                                                 |
| 927.             | H3111B03-5 |                | UNKNOWN H3111B03                                                                                                                               | H3111B03      | Data not found | No Chromosome location info available |                                                                                                                                                                                 |

[0143] The following Examples are intended to illustrate, but not limit, the invention.

### EXAMPLES

#### **Example 1: Signature patterns of gene expression in mouse atherosclerosis and their correlation to human coronary disease**

[0144] Mouse genetic models of atherosclerosis allow systematic analysis of gene expression, and provide a good representation of the human disease process (Breslow (1996) *Science* 272: 685-688). ApoE-deficient mice predictably develop spontaneous atherosclerotic plaques with numerous features similar to human lesions (Nakashima et al. (1994) *Arterioscler Thromb* 14: 133-140; Napoli et al. (2000) *Nutr Metab Cardiovasc Dis* 10: 209-215; Reddick et al. (1994) *Arterioscler Thromb* 14: 141-147. On a high-fat diet, the rate and extent of progression of lesions are accelerated. In addition to environmental influences such as diet, the genetic background of mice has also been found to have an important role in disease development and progression. Whereas C57Bl/6 (C57) mice are susceptible to developing atherosclerosis, the C3H/HeJ (C3H) strain of mice is resistant (Grimsditch et al. (2000) *Atherosclerosis* 151:389-397. Previously, genetic-based diet and age induced transcriptional differences have been demonstrated between these two strains (Tabibiazar et L. (2005) *Arterioscler Thromb Vasc Biol* 25:302-308.

[0145] To more fully characterize the vascular wall gene expression patterns that are associated with atherosclerosis, a systematic large scale transcriptional profiling study was undertaken to take advantage of a longitudinal experimental design, and mouse genetic model and diet combinations that provide varying susceptibility to atherosclerosis. In this experiment, atherosclerosis-associated genes were studied independent of other variables. Primarily, these studies investigated differential gene expression over time in apoE-deficient mice on an atherogenic diet, with comparison to apoE-deficient mice (C57BL/6J-*ApoE*<sup>tm1Unc</sup>) on normal diet as well as C57Bl/6 and C3H/HeJ mice on both normal chow and atherogenic diet. Identification of atherosclerosis-associated genes was facilitated by development of permutation-based statistical tools for microarray analysis which takes advantage of the statistical power of time-course experimental design and multiple biological and technical

replicates. Using these tools, hundreds of known and novel genes that are involved in all stages of atherosclerotic plaque, from fatty streak to end stage lesions, were identified. To further examine the expression of individual genes in the context of particular biological or molecular pathways, a pathway enrichment methodology with gene ontology (GO) terms for functional annotation was utilized. Using classification algorithms, a signature pattern of expression for a core group of mouse atherosclerosis genes was identified, and the significance of these classifier genes was validated with additional mouse and human atherosclerosis samples. These studies identified atherosclerosis related genes and molecular pathways.

## ***Methods***

### *Atherosclerotic lesion analysis*

[0146] For select time points for various experimental groups, 5 to 7 female mice were used for histological lesion analysis. Atherosclerosis lesion area was determined as described previously (Tabibiazar et al. (2005), *supra*). Briefly, the arterial tree was perfused with PBS (pH 7.3) and then perfusion-fixed with phosphate-buffered paraformaldehyde (3%, pH 7.3). The heart and full length of the aorta to iliac bifurcation was exposed and dissected carefully from any surrounding tissues. Aortas were then opened along the ventral midline and dissected free of the animal and pinned out flat, intimal side up, onto black wax. Aortic images were captured with a Polaroid digital camera (DMC1) mounted on a Leica MZ6 stereo microscope, and analyzed using Fovea Pro (Reindeer Graphics, Inc. P. O. Box 2281, Asheville, NC 28802). Percent lesion area was calculated as total lesion area/total surface area.

### *Experimental design, RNA preparation and hybridization to microarrays*

[0147] All experiments were performed following Stanford University animal care guidelines (Saadeddin et al. (2002) *Med Sci Monit* 8:RA5-12). Three week old female apoE knock-out mice (C57BL/6J-*ApoE*<sup>tm1Unc</sup>), C57BL/6J, and C3H/HeJ mice were purchased from Jackson Labs (Bar Harbor, ME). At four weeks of age the mice were either continued on normal chow or were fed high fat diet which included 21% anhydrous milkfat and 0.15% cholesterol (Dyets #101511, Dyets Inc., Bethlehem, PA) for maximum period of 40 weeks. At each of the time-points, including 0 (baseline), 4, 10, 24 and 40 weeks, for each of the

conditions (strain-diet combination), 15 mice (3 pools of 5) were harvested for RNA isolation (total of 405 mice). Additional mice were used for histology for quantification of atherosclerotic lesions as described above. A separate cohort of sixteen-week-old apoE-deficient mice on high fat diet for two weeks (4 pools of 3 aortas) was also used for classification purposes.

[0148] After perfusion of mice with saline, the aortas were carefully dissected in their entirety from the aortic root to the common iliac and subsequently were flash frozen in liquid nitrogen. Total RNA was isolated as described previously (Tabibiazar et al. (2003) *Circ Res* 93:1192-1201) using a modified two-step purification protocol. RNA integrity was also assessed using the Agilent 2100 Bioanalyzer System with RNA 6000 Pico LabChip Kit (Agilent).

[0149] First strand cDNA was synthesized from 10 µg of total RNA from each pool and from a whole 17.5-day embryo for reference RNA in the presence of Cy5 or Cy3 dCTP, respectively. Hybridization to a mouse 60mer oligo microarray (G4120A, Agilent Technologies, Palo Alto, CA) (Carter et al. (2003) *Genome Res* 13:1011-1021) was performed following manufacture's instructions, generating three biological replicates for each of the time points. The RNA from the group of sixteen-week-old mice was linearly amplified and hybridized to a different array (G4121A, Agilent Technologies). Technical validation of the microarray has been performed previously using quantitative real-time reverse transcriptase polymerase chain reaction (results reported in Tabibiazar et al. (2005), *supra*). Primers and probes for 10 representative differentially expressed genes were obtained from Applied Biosystems Assays-on-Demand. A total of 90 reactions, including triplicate assays on three pools of five aortas, was performed from representative RNA samples used for microarray experiments, demonstrating a high correlation between the two platforms (Pearson correlation of 0.82).

#### *Data processing*

[0150] Image acquisition of the mouse oligo microarrays was performed on an Agilent G2565AA Microarray Scanner System and feature extraction was performed with Agilent feature extraction software (version A.6.1.1, Agilent Technologies). Normalization was carried out using a LOWESS algorithm. Dye-normalized signals of Cy3 and Cy5 channels

were used in calculating log ratios. Features with reference values of  $<2.5$  standard deviation for the negative control features were regarded as missing values. Those features with values in at least 2/3 of the experiments and present in at least one of the replicates were retained for further analysis. Reproducibility of microarray results, as measured by the variation between arrays for signal intensities, was assessed using box plots (GeneData, Inc., South San Francisco, CA). For further statistical analysis of the data, a K-nearest-neighbor (KNN) algorithm was applied to impute missing values (Troyanskaya et al. (2001) *Bioinformatics* 17:520-525). Numerical raw data were then migrated into an Oracle relational database (CoBi) that has been designed specifically for microarray data analysis (GeneData, Inc.). Heat maps were generated using "HeatMap Builder" software (Blake and Ridker (2002) *J Intern Med* 252:283-294). All microarray data were submitted to the National Center for Biotechnology information's Gene Expression Omnibus (GEO GSE1560; [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)).

#### *Data analysis*

##### *i) Principal components analysis*

[0151] For each gene the average log expression values were computed at the four post-baseline observation times, 4, 10, 24, and 40 weeks. This was done separately for the six different (diet, strain) combinations, for example ApoE on high fat, presumably the most atherogenic combination. Differences of these vectors were taken for various interesting contrasts, e.g., for ApoE, high-fat minus C3H, normal chow, giving  $N=20280$  vectors of length 4, one for each gene. Principal components analysis of the  $N$  vectors showed a consistent pattern, with the first principal vector indicating a roughly linear increase with observation time.

##### *ii) Time course regression analysis*

[0152] A standard ANACOVA model was fit separately to the log expression values for each gene, using a model incorporating strain, diet, and time period effects. A single important "z value" was extracted from each ANACOVA analysis, for example corresponding to the significance of the time slope difference between the ApoE, high-fat combination and the average of the other five combinations. The  $N$  z-values were then analyzed simultaneously,

using empirical Bayes false discovery rate methods described previously (Efron (2004) *J Amer Stat Assoc* 99:82-95; Efron and Tibshirani (2002) *Genetic Epidemiology* 23:70-86; Efron et al. (2001) *J Amer Stat Assoc* 96:1151-1160). These analyses identified a set of several hundred genes clearly associated with atherosclerosis progression.

*iii) Time course area under the curve analysis*

**[0153]** Area under the curve (AUC) analysis was employed as described previously (Tabibiazar et al. (2005), *supra*). For each sequence of 4 triplicate gene expression measurements over time, the measurement at time 0 was subtracted from all values. The signed area under the curve was then computed. The area is a natural measure of change over time. These areas were then used to compute an F-statistic for the 6 groups (3 mouse strains and 2 diets) and 3 replicates (between sum of squares/within sum of squares). A permutation analysis, similar to that employed in Significance Analysis of Microarrays (SAM) (Tusher et al. *Proc Natl Acad Sci* 98:5116-5121), was carried out to estimate the false discovery rate (q-value or “FDR”) for different levels of the F-statistic.

*iv) Enrichment analysis*

**[0154]** For enrichment analysis, the Expressionist software (GeneData, Inc.), which employs the Fisher exact test to derive biological themes within particular gene sets defined by functional annotation with Gene Ontology (GO) terms ([www.geneontology.org](http://www.geneontology.org)) and Biocarta pathways ([www.biocarta.com/genes/allpathways.asp](http://www.biocarta.com/genes/allpathways.asp)), was used. In this way, over-representation of a particular annotation term corresponding to a group of genes was quantified.

*v) Support vector machine for gene selection*

**[0155]** For supervised analyses, the Expressionist software (GeneData USA), which employs Support Vector Machine (SVM) algorithm (Burges (1998) *Data Mining and Knowledge Discovery* 2:121-167), was used to rank genes based on their utility for class discrimination between time points 0, 4, 10, 24, and 40 weeks in apoE mice on high-fat diet. SVM is a binary classifier, so in order to classify multiple categories, N classifiers were created that classify one group vs. a combination of the rest of the groups (“one vs. all” classifiers) (Ramaswamy et al. (2001) *Proc Natl Acad Sci* 98:15149-15154). The larger set of

genes identified by the time-course analysis was used for this analysis. This method was then used to determine the optimal number of ranked genes to classify the experiments into their correct groups at minimal error rate. The optimal error rate or misclassification is calculated by cross-validation with 25% of the experiments as the test group and the rest as the training group. This is reiterated 1000 times (**Fig. 5A**). In this study, a linear Kernel was used, since a nonlinear Gaussian kernel yielded similar results. This minimal subset of classifier genes was then used for cross-validation as well as classification of other independent gene expression profiling datasets.

*vi) Analysis of independent datasets.*

**[0156]** The SVM algorithm was utilized for classification of independent groups of experiments (Yeang et al. (2001) *Bioinformatics* 17 Suppl 1:S316-322). In this analysis, the primary time-course experiments were used (corresponding to 5 time points mentioned above) as the training set and the independent set of experiments (different array and labeling methodology) as the test set. SVM output for each experiment based on one-versus-all comparisons was represented graphically in a heatmap format (**Fig. 5B**), which is the normalized margin value for each of the 5 SVM classifiers mentioned above. The SVM output permits classification of a new experiment according to the 5 SVM hyperplane. The SVM algorithm (Linear Kernel) was also utilized for external validation by classifying different sets of human expression data. In these analyses, a confusion matrix was generated using cross validation with repeated splits into 75% training and 25% test sets to determine the accuracy of classification based on the small subset of genes identified earlier. Results are represented in tabular fashion (**Table 3**).

*Transcriptional profiling of human atherosclerotic tissue and atherectomy samples*

**[0157]** For one set of samples, coronary arteries were dissected from explanted hearts of patients undergoing orthotopic heart transplantation. Arteries were divided into 1.5 cm segments, classified as lesion or non-lesion after inspection of the luminal surface under a dissecting microscope. RNA was isolated from each individual sample and hybridized to a microarray. A central portion (1-2mm) of each segment was removed and stored in OCT for later histological staining (hematoxylin and eosin, Masson's trichrome). Samples (n=40) were

derived from 17 patients (male 13, female 4, mean age 43 years). Six patients had a diagnosis of ischemic cardiomyopathy, while 11 were classified as non-ischemic, although some vessel segments from the latter had microscopic evidence of coronary artery disease. Of 21 diseased segments, 7 were classified as grade I, 4 grade III and 9 grade V, according to the modified American Heart Association criteria (Virmani et al. (2000) *Arterioscler Thromb Vasc Biol* 20:1262-1275), and one sample had only macroscopic information available. For a second set of tissues, coronary atherectomy samples were obtained with a cutting atherectomy catheter system (Fox Hollow Inc., Redwood City, CA), for chronic atherosclerosis lesions (n=28) and in-stent restenosis lesions (n=14). Patient characteristics in both groups were similar (male 78% vs. 71%, mean age 64 vs. 67). RNA was isolated from each individual sample, labeled by direct or linear amplification methods, and hybridized as described above to a 22k feature custom cardiovascular oligonucleotide microarray designed in conjunction with Agilent Technologies (G2509A, Agilent Inc., Palo Alto, CA). Common reference RNA for all human hybridizations was a mixture of 80% HeLa cell RNA and 20% human umbilical vein endothelial cell RNA. Data processing and analysis were performed as described above. For 2-class comparison of gene expression, Significance Analysis of Microarrays (SAM) was used ([www-stat.stanford.edu/tibs/SAM/](http://www-stat.stanford.edu/tibs/SAM/); Tibibiazar et al. (2003), supra; Tusher et al. (2002), supra).

## ***Results and Discussion***

### *Atherosclerosis in the genetic models*

[0158] To correlate the gene expression results with the extent of disease in each experimental group, the total atherosclerotic plaque burden in the aorta was determined by calculating a percent lesion area from the ratio of atherosclerotic area to total surface area. ApoE-deficient mice (C57BL/6J-*ApoE*<sup>tm1Unc</sup>) (n=7) on high-fat diet were compared to other control mice (n=5-7 for each mouse-diet combination). Representative time-intervals were used for analysis, including baseline measurements in mice prior to initiation of high-fat diet at 4 weeks and end-point measurements corresponding to 40 weeks on either high-fat or normal diet (**Figs. 1, 2**). Gross histological evaluation of these mice demonstrated increased atherosclerotic lesions in ApoE-deficient mice on high-fat diet involving about 50% of the entire aorta, and lesser area involved in ApoE-deficient mice on normal diet (**Fig. 2**). As

expected, the control mice on either diet did not demonstrate evidence of atherosclerosis throughout the course of the experiment (Jawien et al. (2004) *J Physiol Pharmacol* 55:503-517; Nishina et al. (1990) *J Lipid Res* 31:859-869). Although some fatty infiltrates were noted on histological evaluation of the aortic root in C57 mice on high-fat diet, there were no obvious changes in inflammatory cell infiltrate (Tabibiazar et al. (2005), *supra*). The metabolic and lipid profiles of these mice were not obtained in this study, since they are well described in the literature (Grimsditch et al., *supra* ; Nishina et al. (1990), *supra*; Nishina et al. (1993) *Lipids* 28:599-605).

#### *Temporal patterns of gene expression*

**[0159]** Employing a number of mouse models with different propensity to develop atherosclerosis, two different diets, and a longitudinal experimental design, it was possible to factor out differentially regulated genes that are unlikely to be related to the vascular disease process in the apoE deficient model. For instance, age-related and diet-related gene expression patterns that are not linked to vascular disease were eliminated by virtue of their expression in the genetic models that did not develop atherosclerosis. However, the complexity of the experimental design provided significant difficulties related to statistical analysis. Although analytic methods have been proposed to address a single set of time-course microarray data (Luan and Li (2003) *Bioinformatics* 19:474-482; Park et al. (2003) *Bioinformatics* 19:694-703; Peddada et al. (2003) *Bioinformatics* 19:834-841; Xu and Li (2003) *Bioinformatics* 19:1284-1289), there was no accepted algorithm for comparing differences in patterns of gene expression across multiple longitudinal datasets.

**[0160]** Using principle component analysis, it was determined that the greatest variation in the data was between time points, correlating with the progression of disease described previously for the apoE knockout mouse on high fat diet (Nakashima et al. (1994) *Arterioscler Thromb* 14:133-140; Reddick et al. (1994) *Arterioscler Thromb* 14:141-147). Given this finding, a linear regression model was utilized to identify genes that were differentially expressed in ApoE-deficient mice on high-fat diet, compared with all other experimental groups across time. This comparison across strains and dietary groups was employed to focus the analysis on atherosclerosis-specific genes, taking into account gene expression changes in the vessel wall associated with aging, diet, and genetic background. Empirical Bayes and

permutation methods were employed to derive a false discovery rate (FDR) and minimize false detection due to multiple testing. With high stringency limits, global FDR <0.05 and local FDR <0.3, 667 genes demonstrated a linear increase with time, whereas only 64 genes showed the opposite profile (**Fig. 3**).

*Genes with increased expression in the atherosclerotic vessel wall*

[0161] The identification of known genes previously linked to atherosclerosis validated the methodology and analysis algorithm. Most striking in this regard were inflammatory genes, including chemokines and chemokine receptors, such as *Ccl2*, *Ccl9*, *CCr2*, *CCr5*, *Cklfsf7*, *Cxcl1*, *Cxcl12*, *Cxcl16*, and *Cxcr4* (**Fig. 3**). Also upregulated were interleukin receptor genes, including *IL1r*, *IL2rg*, *IL4ra*, *IL7r*, *IL10ra*, *IL13ra*, and *IL15ra*, and major histocompatibility complex (MHC) molecules such as *H2-EB1* and *H2-Ab*. The value of transcriptional profiling in this disease was demonstrated by the identification of numerous inflammatory genes not previously linked to atherosclerosis, including *CD38*, *Fcer1g*, *oncostatin M* (*Osm*) and its receptor (*Osmr*).

[0162] Oncostatin M (*Osm*) and its cognate receptor (*Osmr*) are likely to have significant roles in atherosclerosis, based on number of studies that suggest several important related functions for these genes (Mirshahi et al. (2002) *Blood Coagul Fibrinolysis* 13:449-455. *Osm* is a member of a cytokine family that regulates production of other cytokines by endothelial cells, including *Il6*, *G-CSF* and *GM-CSF*. *Osm* also induces *Mmp3* and *Timp3* gene expression via JAK/STAT signaling (Li et al. (2001) *J Immunol* 166:3491-3498). It induces cyclooxygenase-2 expression in human vascular smooth muscle cells (Bernard et al. (1999) *Circ Res* 85:1124-1131), as well as *Abca1* in HepG2 cells (Langmann et al. (2002) *J Biol Chem* 277:14443-14450). Interestingly, *Stat1*, *Jak3*, *Cox2*, and *Abca1* were among the disease-associated upregulated genes. Additionally, *Osm* produced by macrophages may contribute to development of vascular calcification (Shioi et al. (2002) *Circ Res* 91:9-16). This may occur via regulation of osteopontin or osteoprotegerin (Palmqvist et al. (2002) *J Immunol* 169:3353-3362, both of which have demonstrated significant changes in the dataset described herein. Osteopontin (*Spp1*) is thought to mediate type-1 immune responses (Ashkar et al. (2000) *Science* 287:860-864. While *Spp1* has been extensively studied in atherosclerosis and other immune diseases, some of the osteopontin-related genes identified through these studies

are novel and provide additional links between inflammation and calcification. Some of these include *Cd44*, *Hgf*, osteoprotegerin, *Mglap*, *Il10ra*, *Infgr*, *Runx2*, and *Ccnd1*. *Ibsp*, (*sialoprotein II*), was also noted to be upregulated in these studies. Despite its similar expression profile to *Spp1* in various cancer types and its binding to the same *alpha-v/beta-3 integrin*, the role of *Ibsp* in atherosclerosis has not been elucidated.

[0163] Known and novel genes were identified for many other protein classes that have been studied in atherosclerosis. Genes encoding endothelial cell adhesion molecules were among these groups, including *Alcam* and *Vcam1*. Extracellular matrix and matrix remodeling proteins were found to be upregulated, including *fibronectin*, *Col8a1*, *Ibsp*, *Igsf4*, *Itga6*, and *thrombospondin-1*. Matrix metalloproteinase genes such as *Mmp2* and *Mmp14* as well as those encoding tissue inhibitors of metalloproteinases, including *Timp1*, were also among the upregulated genes. Many transcription factors, lipid metabolism and vascular calcification genes, as well as macrophage and smooth muscle cell specific genes, were among those found to be upregulated. New genes were identified in each of these classes, for example, members of the ATP-binding-cassette family that were not previously associated with atherosclerosis were identified through these studies, including *Abcc3* and *Abcb1b*.

[0164] Interesting genes linked to atherosclerosis for the first time through these studies encode a variety of functional classes of proteins. For example, genes encoding transcription factors *Runx2* and *Runx3* were linked to atherosclerosis in these studies. Cytoplasmic signaling molecules *Vav1*, *Hras1*, and *Kras2* are factors that are well known to have critical signaling functions, but their role in atherosclerosis has not yet been defined. *Wisp1* is a secreted wnt-stimulated cysteine-rich protein that is a member of a family of factors with oncogenic and angiogenic activity. *Rgs10* is a member of a family of cytoplasmic factors that regulate signaling through Toll-like receptors and chemokine receptors in immune cells. Among the new classes of genes identified through these studies to be upregulated in atherosclerosis were those encoding histone deacetylases. Among those genes identified were *Hdac7* and *Hdac2*. Although there is significant evidence that HDACs have important functions regulating growth, differentiation and inflammation, these molecules have not been well studied in the context of atherosclerosis (Dressel et al. (2001) J Biol Chem 276:17007-17013); Ito et al. (2002) Proc Natl Acad Sci 99:8921-8926). Histone deacetylase inhibitors

have been postulated to modulate inflammatory responses (Suuronen et al. (2003) *Neurochem* 87:407-416).

[0165] The data from the experiments described herein has also yielded numerous ESTs and uncharacterized genes. These genes may be attractive candidates for further characterization. One example of such ESTs is *2510004L01Rik*, a gene termed “viral hemorrhagic septicemia virus induced gene” (*VHSV*), which was originally cloned from interferon-stimulated macrophages. This gene is enriched in bone marrow macrophages, is upregulated by CMV infection and is similar to human inflammatory response protein 6 (Chin and Cresswell (2001) *Proc Natl Acad Sci* 98:15125-15130). Several ESTs such as *5930412E23Rik* and *2700094L05Rik* have been cloned from hematopoietic stem cells (genome-www5.stanford.edu/cgi-bin/source/sourceSearch), consistent with data suggesting cells in the diseased vessel wall may emanate from the bone marrow (Rauscher et al. (2003) *Circulation* 108:457-463).

*Genes with decreased expression in the atherosclerotic vessel wall*

[0166] The 64 genes that showed decreased expression during progression of atherosclerosis were of interest, given the lack of previous attention to such genes. *Sparcl1* (*Hevin*) is an extracellular matrix protein which is downregulated in the dataset described herein, and may have antiadhesive (Girard and Springer (1996) *J Biol Chem* 271:4511-4517) and antiproliferative (Claeskens et al. (2000) *Br J Cancer* 82:1123-1130) properties. It has been shown to be downregulated in neointimal formation and suggested to have a possible protective effect in the vessel wall (Geary et al. (2002) *Arterioscler Thromb Vasc Biol* 22:2010-2016). Another gene with decreased expression, *Tgfb3*, may also have a protective effect. The factor encoded by this gene has been shown to decrease scar formation, and to exert an inhibitory effect on G-CSF, suggesting an anti-inflammatory role that would counter pro-inflammatory factors in the vascular wall (Hosokawa et al. (2003) *J Dent Res* 82:558-564); Jacobsen et al. (1993) *J Immunol* 151:4534-4544).

[0167] Interestingly, numerous genes characteristic of various muscle lineages were shown to be downregulated. For smooth muscle cells, this might reflect decreased expression of differentiation markers. For example, the smooth muscle cell gene *caldesmon* encodes a marker of differentiated smooth muscle cells (Sobue et al. (1999) *Mol Cell Biochem* 190:105-

118), and previous studies have noted that the population of differentiated contractile smooth muscle cells that express caldesmon is relatively lower in atherosclerotic plaque (Glukhova et al. (1988) *Proc Natl Acad Sci* 85:9542-9546). Other potential smooth muscle cell marker genes with decreased expression included *Csrp1* and *Mylk*. Other downregulated skeletal and cardiac muscle genes included calsequesterin, which is expressed in fast-twitch skeletal muscle, *Usmg4*, which is upregulated during skeletal muscle growth, *Xin*, which is related to cardiac and skeletal muscle development, and *Sgcg*, that is strongly expressed in skeletal and heart muscle as well as proliferating myoblasts. The possible association of these and other myocyte related genes identified in this study to normal vascular function is not known.

#### *Pathways analysis*

**[0168]** To identify important biological themes represented by genes differentially expressed in the atherosclerotic lesions, the genes were functionally annotated using Gene Ontology (GO) terms ([www.geneontology.org](http://www.geneontology.org)) and curated pathway information. Enrichment analysis with the Fisher Exact Test demonstrated several statistically significant ontologies (**Table 3**), including several associated with inflammation. Inflammatory processes such as immune response, chemotaxis, defense response, antigen processing, inflammatory response, as well as molecular functions such as interleukin receptor activity, cytokine activity, cytokine binding, chemokine and chemokine receptor activity, *Tnf*-receptor, and *MHC I* and *II* receptor activity were noted to be significantly over-represented in the group of genes upregulated with atherosclerosis. Subanalysis of the inflammatory response pathways revealed genes characteristic of the macrophage lineage, as well as both the TH-1 and TH-2 T-cell populations, to be over-represented. Biocarta terms further delineated novel genes that were associated with pathways within the inflammation category, including classical complement, *Rac-CyclinD*, *Egf*, and *Mrp* pathways, as well as those known to be differentially regulated in atherosclerosis, such as *Il2*, *Il7*, *Il22*, *Cxcr4*, *CCr3*, *Ccr5*, *Fcer1*, and *Infg* pathways.

**[0169]** In addition to inflammation, other biological processes and molecular functions were over-represented in the group of differentially upregulated genes. These included expected pathways such as wound healing, ossification, proteo- and peptidolysis, apoptosis, nitric oxide mediated signal transduction, cell adhesion and migration, and scavenger receptor activity. However, several pathways that are less known for their role in atherosclerosis were

also identified, including carbohydrate metabolism, complement activation, calcium ion hemostasis, collagen catabolism, glycosyl bonds and hydrolase activity, taurine transporter activity, heparin activity, etc. The lack of oxygen radical metabolism among the significant processes was surprising, but consistent with up-regulation of genes related to oxygen radical metabolism in all groups with aging.

[0170] Taken together, these pathway analyses support prior observations regarding the importance of inflammatory molecular pathways in atherosclerosis, but additionally, expand the repertoire of molecular pathways that are involved in this disease process.

*Identification of other time-related patterns of gene expression in atherosclerosis*

[0171] The above analysis examined in detail genes with increased expression levels which correlate with atherosclerotic plaque development. However, additional patterns of gene expression were also identified in these longitudinal studies, to identify classes of genes and pathways not previously identified. For these analyses, the AUC algorithm was employed, which measured expression changes over time, made comparisons between the different strain/diet longitudinal datasets to identify gene expression changes specific for the apoE knockout model, and employed permutation to estimate the FDR (Tabibiazar et al. (2005), *supra*). Using this methodology several distinct gene expression patterns and pathways that reflect particular biological processes were identified (**Fig. 4**). For instance, some disease-related pathways were upregulated very early in the disease process and downregulated thereafter (Pattern 6). Others were upregulated early and maintained at relative high expression throughout the time course of the disease (Pattern 8). Whereas the earlier pattern is enriched in pathways representing biological processes such as extracellular matrix and collagen metabolism, as well as DNA replication and response to stress, the later pattern is enriched in pathways representing biological processes such as fatty acid metabolism, oxidoreductase activity and heat-shock protein activity. Some disease related pathways were upregulated in both early and late phases of disease development (Pattern 3), including those associated with metabolism, such as glycolysis and gluconeogenesis. Other patterns (Pattern 4) are represented by key pathways regulating plaque development, including growth factor, cytokine, and cell adhesion activity. Interestingly, inflammation is represented in almost all of the patterns described herein.

*Identification of stage specific gene expression signature patterns*

[0172] Classification approaches to human cancer have provided significant insights regarding the clinical features of the tumor, including propensity to metastasis, drug responsiveness, and long term prognosis (Golub et al. (1999) *Science* 286:531-537; Lapointe et al. (2004) *Proc Natl Acad Sci* 101:811-816; Paik et al. (2004) *N Engl J Med* ("Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer"); Sorlie et al. (2001) *Proc Natl Acad Sci* 98:10869-10874). For atherosclerosis, the clinical utility of classification algorithms will include prediction of future events. To establish a panel of genes whose expression in the vessel wall can accurately classify disease stage, and which may thus be useful for clinical genomic and biomarker applications, the support vector machines algorithm was employed on this comprehensive mouse model disease data set. Employing the SVM classification algorithm, 38 genes were identified that were able to accurately classify each experiment with one of five defined stages of atherosclerosis in mice (**Fig. 5A**). The results demonstrated that these genes can distinguish normal from severe lesions with 100% accuracy. The intermediate stages of the disease are also distinguished from the other stages with a high degree of accuracy (88- 97%) (**Table 3**).

[0173] To validate the classifier genes, their ability to accurately categorize an independent group of 16 week old apoE knockout mice, which were evaluated with a different array and labeling methodology, was evaluated. The microarray utilized different probes for some of the same genes. Moreover, the labeling methodology used a linear amplification step which may introduce further variability in the data. Using the SVM classification algorithm, each of the 4 replicate experiments was accurately classified with the correct stage of the disease process (**Fig. 5B**). As indicated by the greater correlation between gene expression in this independent group of mice and gene expression patterns in the original experimental group aged 24 weeks, the classifier genes accurately matched this validation dataset to the closest timepoint in the database.

*Identification of mouse disease gene expression patterns in human coronary atherosclerosis*

[0174] The expression profile of differentially regulated mouse genes was investigated in human coronary artery atherosclerosis. For transcriptional profiling of human atherosclerotic

plaque, 40 coronary artery samples, dissected from explanted hearts of 17 patients undergoing orthotopic heart transplantation, were used. Of the 21 diseased segments, lesions ranged in severity from grade I to V (modified American Heart Association criteria based on morphological description (Virmani et al., *supra*)). For the purpose of this analysis, human artery segments were classified as non-lesion or lesion (combined all grades). Atherosclerosis related mouse genes were matched to human orthologs by gene symbol or by known homology ([www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=homologene](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=homologene)). Comparison of expression of the mouse genes between lesion and non-lesion human samples using the significance analysis of microarrays algorithm (FDR <0.025) revealed more than 100 mouse genes with higher expression in the diseased human tissue (**Fig. 6**). In view of the differences between the tissue samples used in these gene expression experiments, these constitute an important common set of disease relevant genes.

[0175] To further test the relevance of our findings in mouse atherosclerosis, the accuracy of the mouse classifier genes was assessed in human atherosclerotic disease, employing established statistical methods. The mouse classifier genes were first used to predict various stages of coronary artery disease in the human arterial samples. The results demonstrated a high degree of accuracy in predicting atherosclerotic disease severity (71.2 to 84.7% accuracy) (**Table 3**).

[0176] Additionally, the mouse classifier genes were used to categorize human atherectomy tissue obtained from coronary vessels treated for chronic atherosclerosis or in-stent restenosis. The pathophysiological basis of restenosis is quite distinct from that of chronic coronary atherosclerosis, and it was of interest to demonstrate that the classifier genes could distinguish the disease processes (Rajagopal and Rockson (2003) *Am J Med* 115:547-553). The results (**Table 3**) demonstrated significant accuracy in distinguishing the two types of lesions (85.4 to 93.7% accuracy), further validating the significance of the mouse atherosclerosis gene expression patterns in human disease. The greater accuracy of classification with these samples compared to the arterial segments likely reflects less variation in the clinical profile of the patients, which have much less complex medication and comorbid features than the pre-cardiac transplant patients in the above analysis.

**Table 2.** Biological themes in atherosclerosis. Enrichment analysis of atherosclerosis-related genes annotated with Gene Ontology and Biocarta terms demonstrates involvement of multiple molecular pathways and biological processes. Probabilities (p-values) were derived using Fisher exact test. 8478 of the entire microarray and 513 of genes in our set (including additional 183 genes which demonstrated Pearson correlation >0.8 with the upregulated pattern) were annotated with GO, Biocarta, or other terms.

| Biological Process (GO annotation)                        | List gene # | Total gene # | p-value | Molecular Function (GO annotation)              | List gene # | Total gene # | p-value | Biocarta Pathways | List gene # | Total gene # | p-value |
|-----------------------------------------------------------|-------------|--------------|---------|-------------------------------------------------|-------------|--------------|---------|-------------------|-------------|--------------|---------|
| immune response                                           | 10          | 73           | <0.0001 | acting on pyrolysis fumes                       | 2           | 1            | <0.0001 | biocartaPathway   | 3           | 3            | <0.0001 |
| cell surface receptor linked signal transduction          | 12          | 21           | <0.0001 | antioxidant activity                            | 5           | 13           | <0.0001 | biocartaPathway   | 4           | 7            | <0.0001 |
| defense response                                          | 15          | 38           | <0.0001 | hydrolase activity                              | 67          | 641          | <0.0001 | actPathway        | 5           | 12           | <0.0001 |
| carbohydrate metabolism                                   | 14          | 60           | <0.0001 | cytokine activity                               | 13          | 57           | <0.0001 | actPathway        | 5           | 13           | <0.0001 |
| anigen processing                                         | 5           | 9            | <0.0001 | hematopoietin                                   | 9           | 32           | <0.0001 | actPathway        | 4           | 8            | <0.0001 |
| locomotory behavior                                       | 4           | 6            | <0.0001 | complement activity                             | 5           | 9            | <0.0001 | actPathway        | 4           | 8            | <0.0001 |
| inflammatory response                                     | 8           | 30           | <0.0001 | cytokine binding                                | 3           | 3            | <0.0001 | actPathway        | 3           | 4            | <0.0001 |
| complement activation                                     | 5           | 12           | <0.0001 | C-C chemokine receptor activity                 | 3           | 3            | <0.0001 | actPathway        | 4           | 10           | <0.0001 |
| proteolysis and peptidolysis                              | 25          | 204          | <0.0001 | chemokine activity                              | 4           | 7            | <0.0001 | actPathway        | 5           | 17           | <0.0001 |
| anigen presentation                                       | 4           | 10           | <0.0001 | cysteine-type endopeptidase activity            | 11          | 63           | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| intracellular signaling cascade                           | 28          | 269          | <0.0001 | tumor necrosis factor receptor activity         | 2           | 2            | <0.0001 | actPathway        | 3           | 7            | <0.0001 |
| zinc ion homeostasis                                      | 2           | 2            | <0.0001 | platelet-derived growth factor receptor binding | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| anion channel activity                                    | 2           | 2            | <0.0001 | cathepsin D activity                            | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| transmembrane receptor protein tyrosine kinase activation | 2           | 2            | <0.0001 | beta-H-acylthioesterase activity                | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| hair cell differentiation                                 | 2           | 2            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| cell death                                                | 2           | 2            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| exogenous antigen via MHC class II                        | 3           | 7            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| osteification                                             | 4           | 14           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| collagen catabolism                                       | 3           | 8            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| classical pathway                                         | 3           | 8            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| vesicle transport along actin filament                    | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| nitric oxide mediated signal transduction                 | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| negative regulation of angiogenesis                       | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| endogenous antigen via MHC class I                        | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| cellular defense response (sensu Vertebrata)              | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| beta-alanine transport                                    | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| lymph gland development                                   | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| perception of pain                                        | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| myeloid blood cell differentiation                        | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| female gamete generation                                  | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| cytolysis                                                 | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| ATP biosynthesis                                          | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| regulation of peptidyl-tyrosine phosphorylation           | 3           | 11           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| neurotransmitter transport                                | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| neurogenesis                                              | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| cell adhesion                                             | 2           | 3            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| regulation of cell migration                              | 3           | 13           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| wound healing                                             | 2           | 6            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| ureteric bud branching                                    | 2           | 6            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| cellular defense response                                 | 2           | 6            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| acute-phase response                                      | 2           | 6            | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| regulation of transcription from Pol II promoter          | 6           | 44           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| hydrogen transport                                        | 3           | 14           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |
| calcium ion homeostasis                                   | 3           | 14           | <0.0001 | autotaxin activity                              | 2           | 2            | <0.0001 | actPathway        | 2           | 2            | <0.0001 |

**Table 3.** Classification of mouse and human atherosclerotic tissues employing mouse classifier genes. To validate the accuracy of mouse classifier genes in predicting disease severity we utilized various mouse and human expression datasets. The SVM algorithm was utilized for cross validation of mouse experiments grouped on the basis of (A) stage of disease (no disease- apoE time 0, mild disease- apoE at 4 and 10 weeks on normal diet, mild-moderate disease- apoE at 4 and 10 weeks on highfat diet, moderate disease-apoE at 24 and 40 weeks on normal diet, and severe disease- apoE at 24 and 40 weeks on high fat diet); (B) 3 different time points (apoE at 0 vs. 10, vs. 40 weeks); (C) Human coronary artery with lesion vs. no lesion; and (D) atherectomy samples derived from in-stent restenosis vs. native atherosclerotic lesions. For each analysis, the accuracy of classification is represented in tabular fashion with the confusion matrix generated using N-fold cross validation methods.

|                       |                     |                     |                     |                |                   |             |
|-----------------------|---------------------|---------------------|---------------------|----------------|-------------------|-------------|
| <b>A</b><br>PREDICTED | TRUE<br>No dz       | TRUE<br>Mild_dz     | TRUE<br>Mild_mod dz | TRUE<br>Mod_dz | TRUE<br>Severe_dz | Correct [%] |
|                       | 64                  | 0                   | 1                   | 0              | 0                 | 98.5        |
|                       | 2                   | 140                 | 0                   | 0              | 0                 | 98.6        |
|                       | 0                   | 0                   | 148                 | 20             | 0                 | 88.1        |
|                       | 0                   | 0                   | 3                   | 149            | 0                 | 98.0        |
|                       | 0                   | 0                   | 0                   | 0              | 173               | 100.0       |
|                       | 97.0                | 100.0               | 97.4                | 88.2           | 100.0             |             |
| <b>B</b><br>PREDICTED | TRUE<br>ApoE_T00_NC | TRUE<br>ApoE_T10_HF | TRUE<br>ApoE_T40_HF | Correct [%]    |                   |             |
|                       | 68                  | 0                   | 0                   | 100            |                   |             |
|                       | 0                   | 56                  | 0                   | 100            |                   |             |
|                       | 0                   | 0                   | 76                  | 100            |                   |             |
|                       | 100                 | 100                 | 100                 |                |                   |             |
| <b>C</b><br>PREDICTED | TRUE<br>Lesion      | TRUE<br>No lesion   | Correct [%]         |                |                   |             |
|                       | 183                 | 33                  | 84.7                |                |                   |             |
|                       | 53                  | 131                 | 71.2                |                |                   |             |
|                       | 77.5                | 79.9                |                     |                |                   |             |
| <b>D</b><br>PREDICTED | TRUE<br>ISR         | TRUE<br>De novo     | Correct [%]         |                |                   |             |
|                       | 345                 | 44                  | 88.7                |                |                   |             |
|                       | 59                  | 652                 | 91.7                |                |                   |             |
|                       | 85.4                | 93.7                |                     |                |                   |             |

## Example 2: Mouse Strain –Specific Differences in Vascular Wall Gene Expression and Their Relationship to Vascular Disease

### Methods

#### RNA preparation and hybridization to the microarray

[0177] Three-week old female C3H/HeJ, C57Bl/6J, and apoE knock-out mice (C57BL/6J-*ApoE<sup>tm1Unc</sup>*) were purchased from Jackson Labs (JAX® Mice and Services, Bar Harbor, ME). At four weeks of age the mice were either continued on normal chow or switched to non-cholate containing high-fat diet which included 21% anhydrous milkfat and 0.15% cholesterol (Dyets #101511, Dyets Inc., Bethlehem, PA) for a maximum period of 40 weeks. At each of the time-points, including 0 (baseline), 4, 10, 24 and 40 weeks, for each of the conditions (strain-diet combination), 15 mice were harvested for RNA isolation, for a total of 450 mice. Following

Stanford University animal care guidelines, the mice were anesthetized with Avertin and perfused with normal saline. The aortas from the root to the common iliacs were carefully dissected, flash frozen in liquid nitrogen, and divided into three pools of five aortas for further RNA isolation. Total RNA was isolated as described in Tabibiazar et al. (2003) *Circ Res* 93:1193-1201. First strand cDNA was synthesized from 10 µg of total RNA from each pool and from whole 17.5-day embryo for reference RNA in the presence of Cy5 or Cy3 dCTP, respectively, and hybridized to a mouse 60mer oligo microarray (G4120A, Agilent Technologies, Palo Alto, CA), generating three biological replicates for each time point.

#### *Data processing*

[0178] Array image acquisition and feature extraction was performed using the Agilent G2565AA Microarray Scanner and feature extraction software version A.6.1.1. Normalization was carried out using a LOWESS algorithm, and Dye-normalized signals were used in calculating log ratios. Features with reference values of <2.5 standard deviations above background for the negative control features were regarded as missing values. Those features with values in at least 2/3 of the experiments and present in at least one of the replicates were retained for further analysis. For SAM analyses, a K-nearest-neighbor (KNN) algorithm was applied to impute for missing values. (Tabibiazar et al. (2003), *supra*.)

#### *Data analysis*

[0179] Experimental design and analysis flow chart is depicted in **Figure 7**. Significance Analysis of Microarrays (SAM) was employed to identify genes with statistically different expression between the C3H and C57 mice at baseline. (Tabibiazar et al. (2003), *supra*; Tusher et al. (2001) *PNAS* 98:5116-5121; Chen et al. (2003) *Circulation* 108:1432-1439.) For partitioning clustering of the genes with K-Means and self-organizing-maps (SOM), we used positive correlation for distance determination and required complete linkage, which uses the greatest distance between genes to ascribe similarity. SOM and K-Means analyses were performed using Expressionist software (GeneData, Inc., USA). Heatmaps were generated using HeatMap Builder. For enrichment analysis we used the EASE analysis software which employs Gene Ontology (GO) annotation and the Fisher's exact test to derive biological themes within particular gene sets. (Hosack et al. (2003) *Genome Biol.* 4:R70.) For time-course study, a new statistical algorithm, the Area-Under-Curve (AUC) analysis was devised. For each sequence of 4 triplicate gene expression measurements over time, we first subtracted the measurement at time 0 from all

values. We then computed the signed area under the curve. The area is a natural measure of change over time. These areas were then used to compute an F-statistic for comparing C57 and C3H mice across the different diets. A permutation analysis, similar to that employed in SAM, was carried out to estimate the false discovery rate (q-value or “FDR”) for different levels of the F-statistic. For ease of presentation, genes which meet our FDR cutoffs will be referred to as “significant” throughout the remainder of the article. All microarray data were submitted to the NCBI Gene Expression Omnibus (GEO GSE1560; <http://www.ncbi.nlm.nih.gov/geo/>).

#### *Aortic lesion analysis*

[0180] For select time points within various experimental groups, 5 to 7 female mice were used for histological lesion analysis. Atherosclerosis lesion area was determined as described in Tangirala et al. (1995) 36:2320-2328.

#### *Quantitative Real-Time Reverse Transcriptase–Polymerase Chain Reaction*

[0181] Primers and probes for 10 representative differentially expressed genes were obtained from Applied Biosystems Assays-on-Demand. A Total of 90 reactions were performed from representative RNA samples used for microarray experiments. These included triplicate assay on three pools of five aortas. cDNA was synthesized and Taqman was performed as described in Tabibiazar et al. (2003), *supra*.

### **Results**

#### *Baseline differences in gene expression patterns between the mouse strains*

[0182] Differences in gene expression levels between the two strains at baseline, before effects of aging or diet become apparent, may identify genes that play a role in determining vascular wall disease susceptibility. To identify such genes SAM was used to compare the vascular wall gene expression of C3H vs. C57 mice at 4 weeks of age, with all animals on normal chow diet. SAM identified 311 genes as being significantly differentially expressed (FDR <0.1 with >1.5 fold difference), and expression patterns of these genes provided a clear partition between C3H and C57 mice (**Fig. 8**). A separate 2-class comparison (SAM, FDR <0.1) between C57 and apoE-deficient mice with a C57Bl/6 genetic background revealed only a few genes, including Apo-E, which were differentially expressed in the 2 groups of mice (data not shown).

[0183] Comparison of C3H and C57 vascular wall gene expression at baseline provided a list of compelling candidate genes which reflected differences in biological processes such as growth,

differentiation, and inflammation as well as molecular functions such as catecholamine synthesis, phosphatase activity, peroxisome function, insulin like growth factor activity, and antigen presentation (**Fig. 8**). These processes were exemplified by higher expression of genes such as *Cdkn1a*, *Pparbp*, protein tyrosine phosphatase-4a2, and *Socs5* in C3H mice, compared with genes such as *ABCC1*, *H2-D1*, *Bat5*, *IGFBP1*, *SCD1*, and *Serpine6b* which demonstrated higher expression in C57 mice. These fundamental baseline gene expression differences may determine disease susceptibility as the mice are exposed to age-related stimuli or dietary challenges.

*Age-related differences in gene expression patterns between the mouse strains*

[0184] To further examine the vascular wall gene expression differences between C57 and C3H mice, an analysis was performed to identify genes differentially expressed in response to aging (**Fig. 9**). Data was collected at five time points over a 40 week period. To identify such genes, we developed the Area Under the Curve (AUC) analysis. The AUC analysis relies on a permutation procedure to reduce the number of potential false positives generated due to multiple testing, but still utilizes the increase in statistical power of time-course experimental design. Comparing C57 vs. C3H time-course differences on normal diet with a rigid cutoff (FDR <0.05) did not identify any genes. However, relaxing the AUC stringency (f-statistic >10, FDR <0.45) allowed a large number of genes (413) to be included for pathway over-representation analysis using GO annotation. Functional annotation and group over-representation analysis (Fisher test p-value <0.02) of the resultant differentially expressed genes revealed differences in a number of biological processes, including growth and development, as well as a number of molecular functions such as cell cycle control, regulation of mitosis, and metabolism (**Fig. 9b**). Some of these processes are exemplified by genes with higher expression in C57 mice, such as *Aoc1* (pro-oxidative stress), *Bub1* (cell cycle check point), Cyclin B2, as well as genes with higher expression in C3H, including *INHBA* and *INHBB*.

[0185] Temporally variable genes identified by AUC analysis were further characterized with K-Means clustering to identify dynamic patterns of expression during the aging process (**Fig. 3c**). Clusters 1, 4, and 9 revealed either higher overall expression or temporally increasing levels of expression in C3H mice compared with C57 mice. In contrast, clusters 2, 6, and 14 revealed the opposite pattern. Of the genes which were noted to be differentially expressed in the two strains during aging, 51 genes were also differentially expressed at baseline, suggesting that baseline differences of certain genes can further be affected with aging.

*Diet-related differences in gene expression patterns between the mouse strains*

[0186] Differential vascular wall response to atherogenic stimuli was determined by comparing temporal gene expression patterns in C57 vs. C3H mice on high-fat diet (**Fig. 10A**). Comparing C57 vs. C3H time-course differences on high-fat diet with a rigid cutoff (FDR < 0.05) identified 35 genes, including *Hgfl* and *Tgfb4*, which were down regulated in C57 on high-fat diet. Additional known genes, as well as a number of ESTs were also identified. Employing a less stringent AUC cutoff allowed identification of a larger number of genes, which could be evaluated with pathway over-representation analysis using GO annotation. At this level of stringency (f-statistic > 10, FDR < 0.35), a total of 650 genes with temporally variable expression were identified. Genes that were also differentially regulated by the aging process (141 of 650 genes) were excluded from further analysis of this group. 38 of the remaining 509 genes were among those differentially expressed at baseline. Functional annotation and group over-representation analysis (Fisher test p-value < 0.02) of these differentially expressed genes revealed differences in biological processes such as catabolism, oxygen reactive species and superoxide metabolism, and proteo- and peptidolysis as well as molecular functions such as fatty acid metabolism, oxidoreductase and methyltransferase activities (**Fig. 10B**). Interestingly, this analysis suggested important differences between the two mouse strains with respect to the activity of the peroxisome, microbody and lysosome. Some of these processes were exemplified by genes with higher expression in C3H mice, such as *Ccs*, *Ephx2*, *Gpx4*, *Prdx6* (anti-oxidants), *Sirt3* (transcriptional repressor), *PPAR $\alpha$* , and *Mcd*, as well as genes with higher expression in C57 mice, such as Lysyl oxidase and *Cdkn1a*. K-means clustering of these genes identified a small number of distinct expression patterns (**Fig. 10C**), with clusters 3 and 9 revealing increased gene expression in C3H mice and clusters 8 and 10 showing the opposite pattern.

*Evaluation of strain-specific differentially regulated genes in the apoE model*

[0187] Using these techniques, a significant number of genes have been identified that are differentially expressed in the atherosclerosis resistant C3H and susceptible C57 mice, some of which are likely involved in atherogenesis and some of which are likely irrelevant to the process. To further select genes most likely to be involved in atherogenesis, expression in apoE-deficient mice fed normal or high-fat diet over a period of 40 weeks was investigated (**Fig. 11**). We utilized SOM analysis to visualize the expression profiles of these subsets of genes throughout the development and progression of atherosclerosis in the ApoE-deficient mice. The analysis revealed

several patterns of gene expression. For example, SOM cluster 8 demonstrated a consistently increasing pattern of expression which correlated with disease progression in the apoE-deficient mice (**Fig. 11**). As evidenced by the pie chart, this cluster is enriched with genes that were identified as more highly expressed in C57 versus C3H mice at baseline (*i.e.*, potentially atherogenic). In contrast, clusters 4, 5, and 6 showed decreasing expression with disease progression. The decreased expression of genes in cluster 4 was somewhat attenuated with high-fat challenge of the ApoE-deficient mice. This cluster is particularly enriched with genes that had revealed a higher expression in C3H mice (*i.e.*, potentially atheroprotective) with atherogenic stimuli and with aging.

[0188] Given C3H resistance and C57 susceptibility to atherosclerosis, as an initial hypothesis it was postulated that genes with higher expression in C3H mice confer resistance, whereas genes with higher expression in C57 mice may have a pro-atherogenic role. With this point of reference, gene clusters were further examined. For example, limiting the list of genes in SOM cluster 8 (genes with increased expression with atherosclerosis) to those that also had higher baseline expression in C57 mice yielded an interesting set of genes that may be atherogenic. This group included inflammation related genes such as *H2-D1*, *Pdgfc*, *Paf*, and *Cd47*. Other compelling genes included *Agpt2*, *Mglap*, *Xdh*, *Th*, and *Ctsc*. Conversely, limiting the list of genes in clusters 4 and 5 to those with higher expression in C3H mice identified a group of genes with potential athero-protective function. Some of those genes included *Ppara $\alpha$* , *Pparbp*, as well as *Ptp4a1*, and *Mcd*.

#### *Lesion analysis in the genetic models*

[0189] To address whether some of the gene expression differences are related to presence of atherosclerotic lesion in C57 mice, the total atherosclerotic burden was determined in the aorta by calculating a percent lesion area in aortas of C57 (n=5) and C3H (n=5) mice. Comparisons were made at time 0 and 40 weeks on normal or high-fat diet. Non-cholate containing high-fat diet was used to prevent caustic effects on the vascular wall. As expected, C57 and C3H mice on either diet did not demonstrate evidence of atherosclerosis throughout the course of the experiment, suggesting that observed gene expression changes cannot be explained by different cellular composition of the vessel wall. Although minimal fatty infiltrates were noted on histological evaluation of the aortic root in C57 mice on high-fat diet, there were no obvious changes in inflammatory cell infiltrate.

*Quantitative RT-PCR validation of expression differences*

[0190] To validate the array results with quantitative RT-PCR and assure that the statistical analyses were identifying truly differentially expressed genes, ten representative genes were assayed by quantitative RT-PCR. Several genes were used from each group of significant genes. There is high degree of correlation between the two methodologies (Pearson correlation of 0.86), validating the results of the microarray analyses.

[0191] Although the foregoing invention has been described in some detail by way of illustration and examples for purposes of clarity of understanding, it will be apparent to those skilled in the art that certain changes and modifications may be practiced without departing from the spirit and scope of the invention. Therefore, the description should not be construed as limiting the scope of the invention.

[0192] All publications, patents and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent or patent application were specifically and individually indicated to be so incorporated by reference.

CLAIMS

We claim:

1. A system for detecting gene expression, comprising at least two isolated polynucleotide molecules, wherein each of said at least two isolated polynucleotide molecules detects an expressed gene product from a gene that is differentially expressed in atherosclerotic disease in a mammal, wherein said gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927.
2. A system for detecting gene expression, comprising at least two isolated polynucleotide sequences, wherein each of said at least two isolated polynucleotide molecules detects an expressed gene product from a gene that is differentially expressed in atherosclerotic disease in a mammal, wherein said gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.
3. A system for detecting gene expression according to claim 1, wherein at least one of said isolated polynucleotide molecules detects an expressed gene product from a gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.
4. A system according to claim 1, wherein the isolated polynucleotide molecules are immobilized on an array.
5. A system according to claim 4, wherein the array is selected from the group consisting of a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, polynucleotide array or a cDNA array, a microtiter plate, a membrane, and a chip.
6. A system according to claim 1, wherein the isolated polynucleotides are selected from the group consisting of synthetic DNA, genomic DNA, cDNA, RNA, or PNA.
7. A kit comprising the system of claim 1.

8. A kit comprising the system of claim 4.
9. A method of monitoring atherosclerotic disease in an individual, comprising detecting the expression level of at least one gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927.
10. The method of claim 9, wherein said at least one gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.
11. The method of claim 9, comprising detecting the expression level of at least two of said genes.
12. The method of claim 11, wherein at least one of said at least two genes is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.
13. The method of claim 9, comprising detecting the expression level of at least ten of said genes.
14. The method of claim 13, wherein at least one of said at least ten genes is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.
15. The method of claim 9, comprising detecting the expression level of at least one hundred of said genes.
16. The method of claim 15, wherein at least one of said at least one hundred genes is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.

17. The method of claim 9, wherein said atherosclerotic disease comprises coronary artery disease.

18. The method of claim 9, wherein said atherosclerotic disease comprises carotid atherosclerosis.

19. The method of claim 9, wherein said atherosclerotic disease comprises peripheral vascular disease.

20. The method of claim 9, wherein said expression level is detected by measuring the RNA level expressed by said one or more genes.

21. The method of claim 20, comprising isolating RNA from said individual prior to detecting the RNA expression level.

22. The method of claim 20, wherein said detection of said RNA expression level comprises amplifying RNA from said individual.

23. The method of claim 22, wherein amplification of RNA comprises polymerase chain reaction (PCR).

24. The method of claim 20, wherein detection of said RNA expression level comprises hybridization of RNA from said individual to a polynucleotide corresponding to said at least one gene selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 1-927.

25. The method of claim 20, wherein said expression level is detected by measuring the protein level expressed by said one or more genes.

26. The method of claim 9, further comprising selecting an appropriate therapy for said atherosclerotic disease.

27. The method of claim 9, comprising detecting the expression of said at least one gene in serum from said individual.

28. The method of claim 20, comprising measuring said RNA level in serum from said individual.

29. The method of claim 25, comprising measuring said protein level in serum from said individual.

30. A method of monitoring atherosclerotic disease in an individual, comprising detecting RNA expressed from at least one gene selected from the group of genes corresponding to at least one polynucleotide sequence depicted in SEQ ID NOs: 1-927.

31. The method of claim 30, wherein said at least one gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.

32. The method of claim 30, comprising measuring said RNA in serum from said individual.

33. A method of monitoring atherosclerotic disease in an individual, comprising detecting protein expressed from at least one gene selected from the group of genes corresponding to at least one polynucleotide sequence depicted in SEQ ID NOs: 1-927.

34. The method of claim 33, wherein said at least one gene is selected from the group of genes corresponding to the polynucleotide sequences depicted in SEQ ID NOs: 8, 14, 26, 32, 50, 64, 83, 99, 142, 154, 159, 161, 177, 181, 200, 390, 430, 434, 439, 440, 476, 491, 508, 530, 534, 565, 567, 572, 624, 647, 657, 690, 733, 745, 806, 824, 886, 882, 901, 905, 913, and 927.

35. The method of claim 33, comprising measuring said protein in serum from said individual.

Figure 1

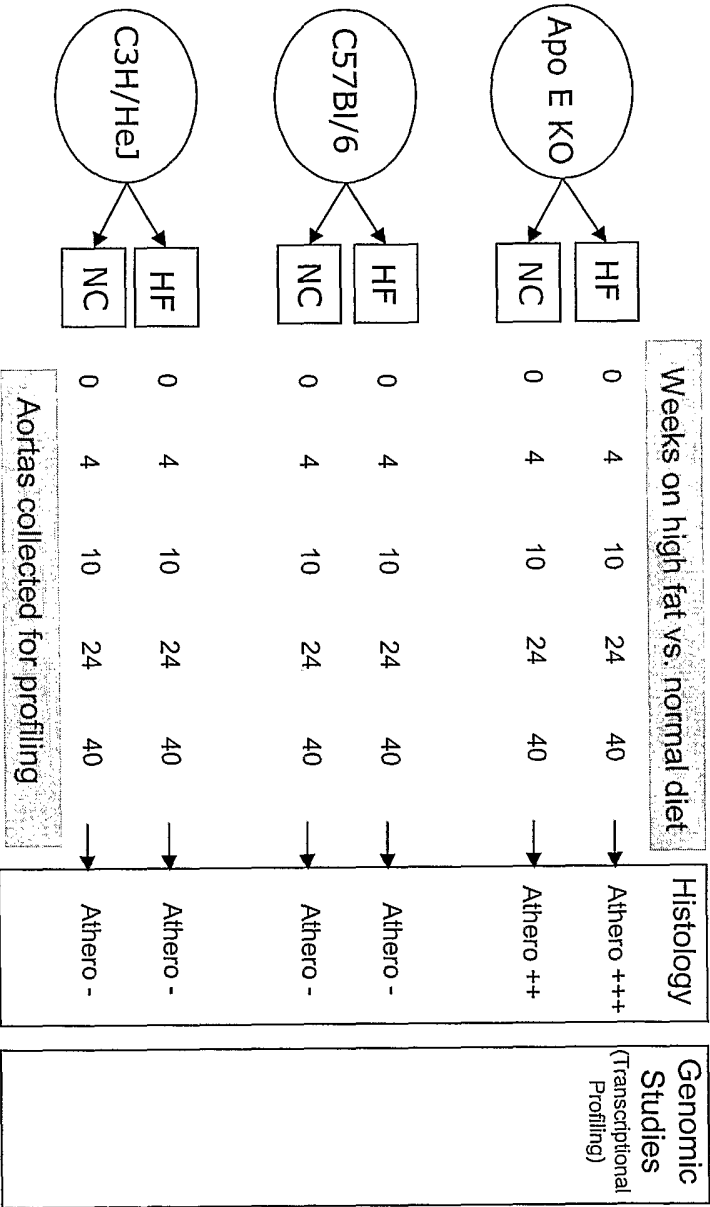
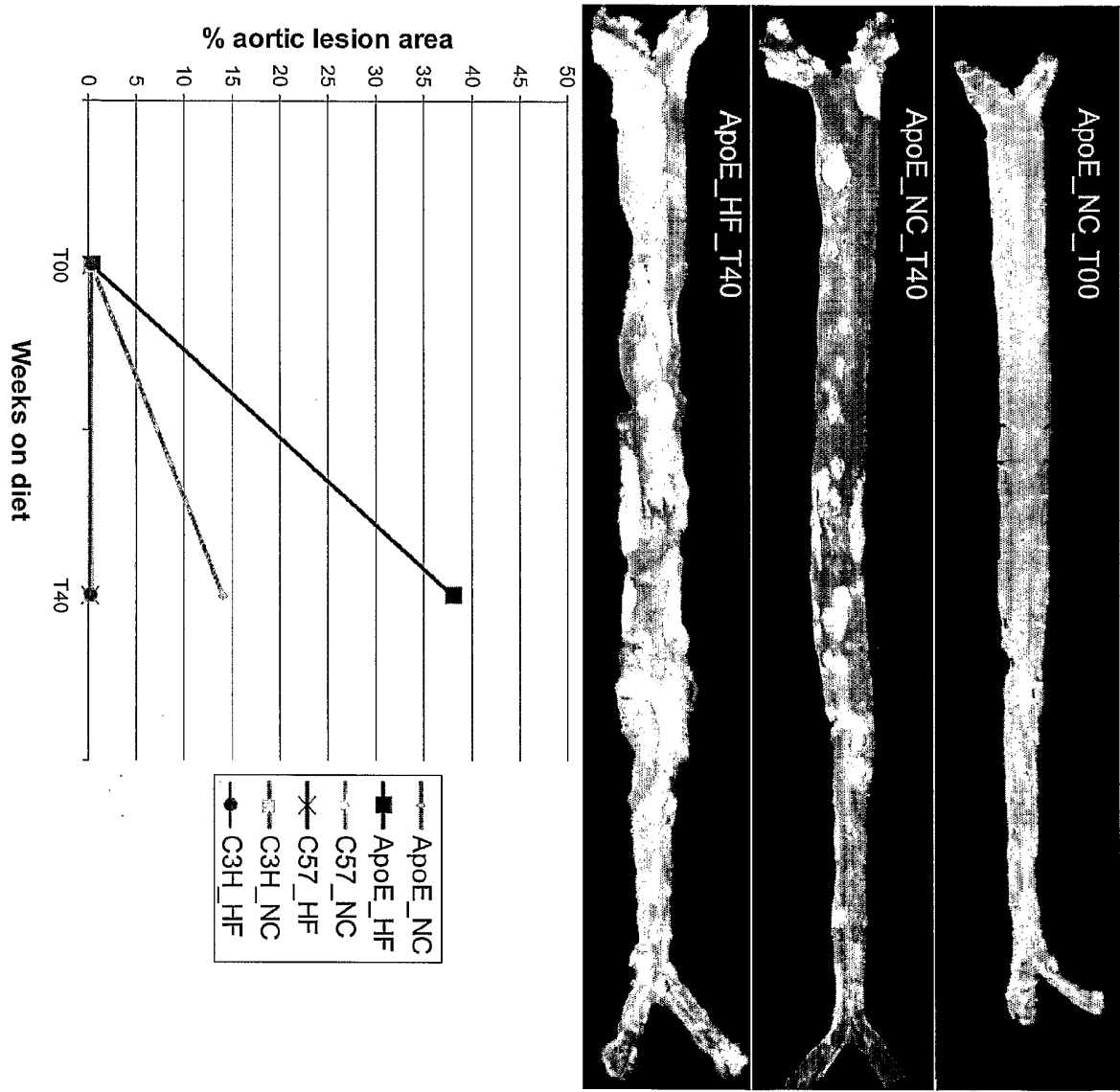


Figure 2



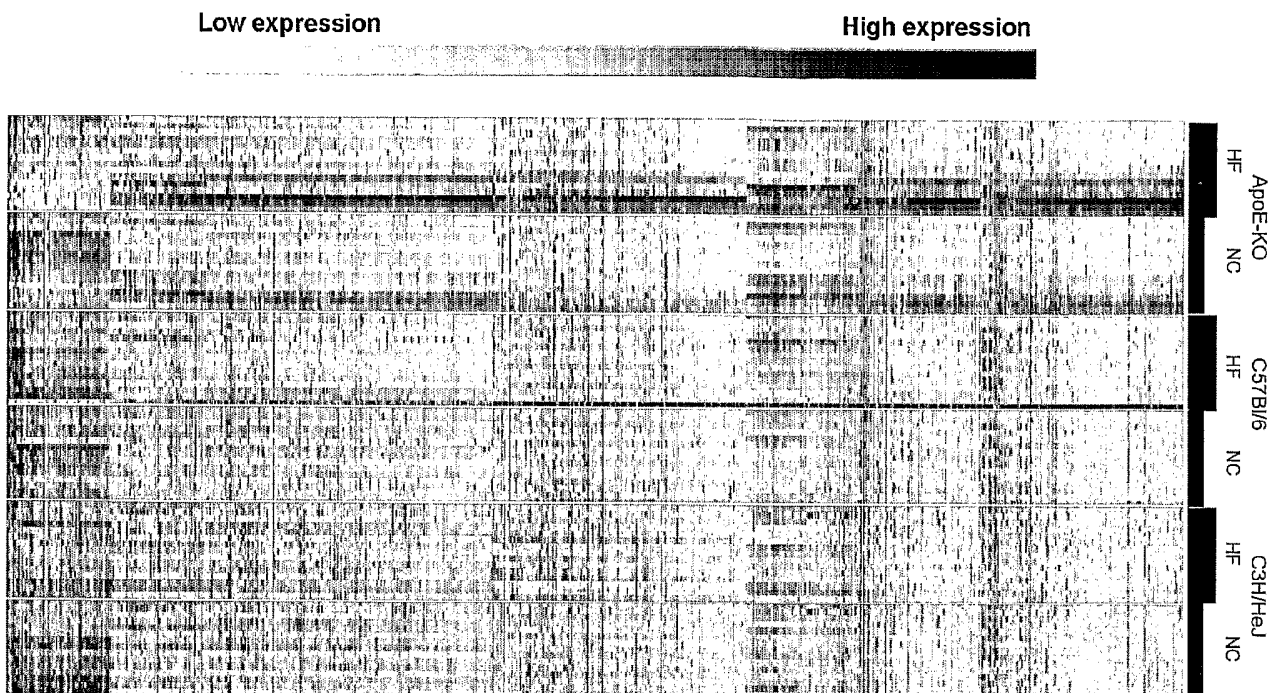


Figure 3

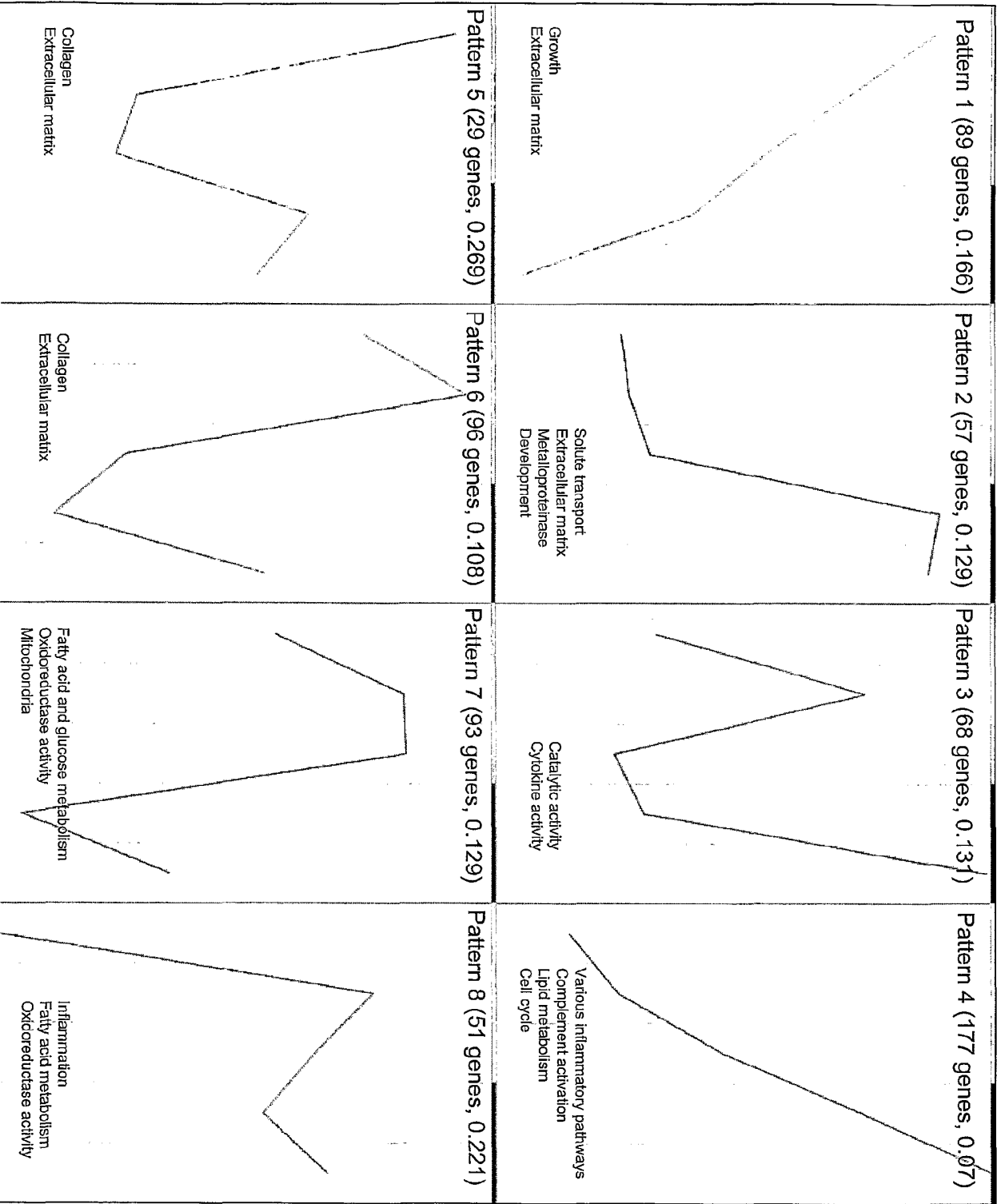


Figure 4

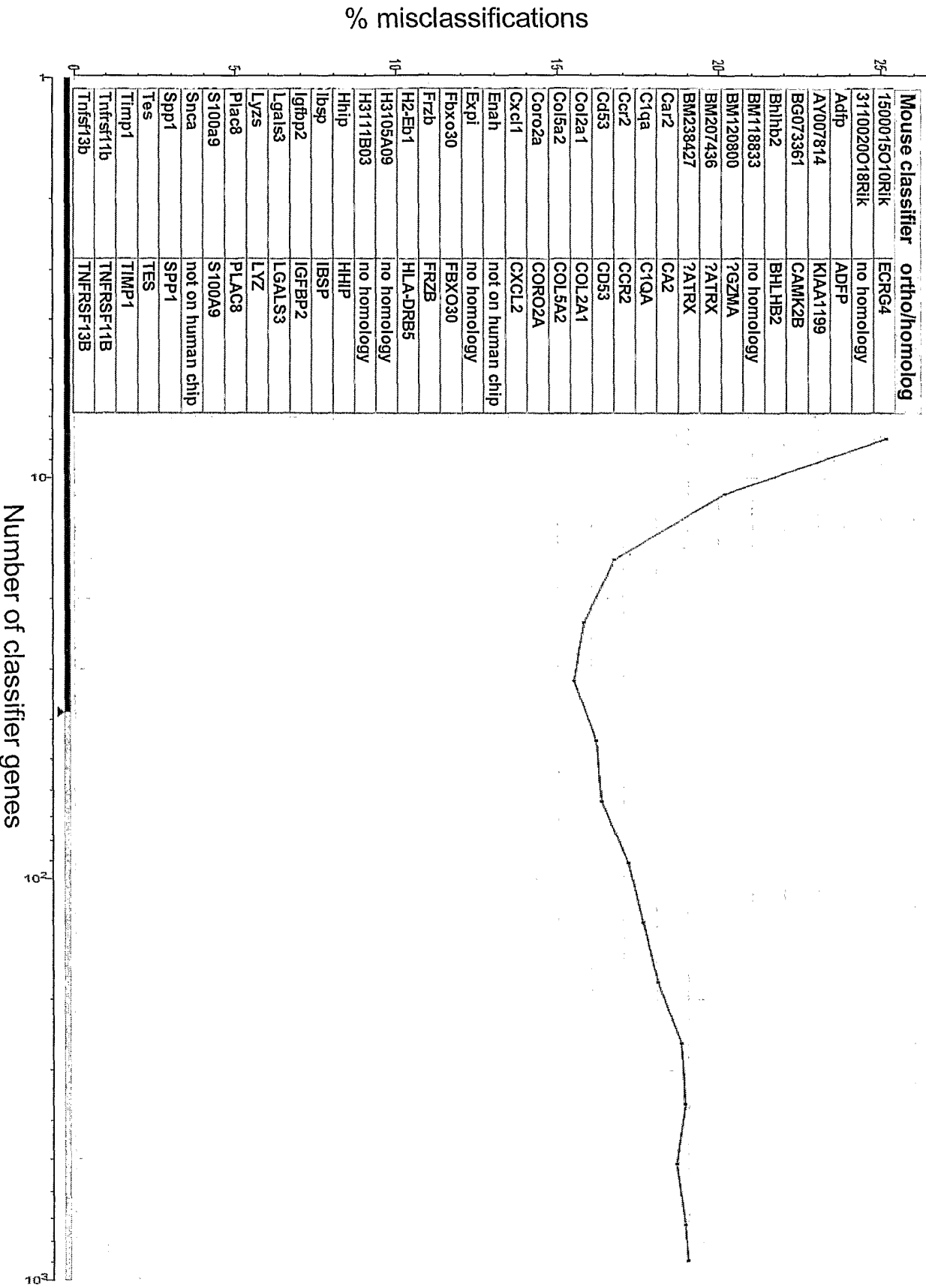


Figure 5A

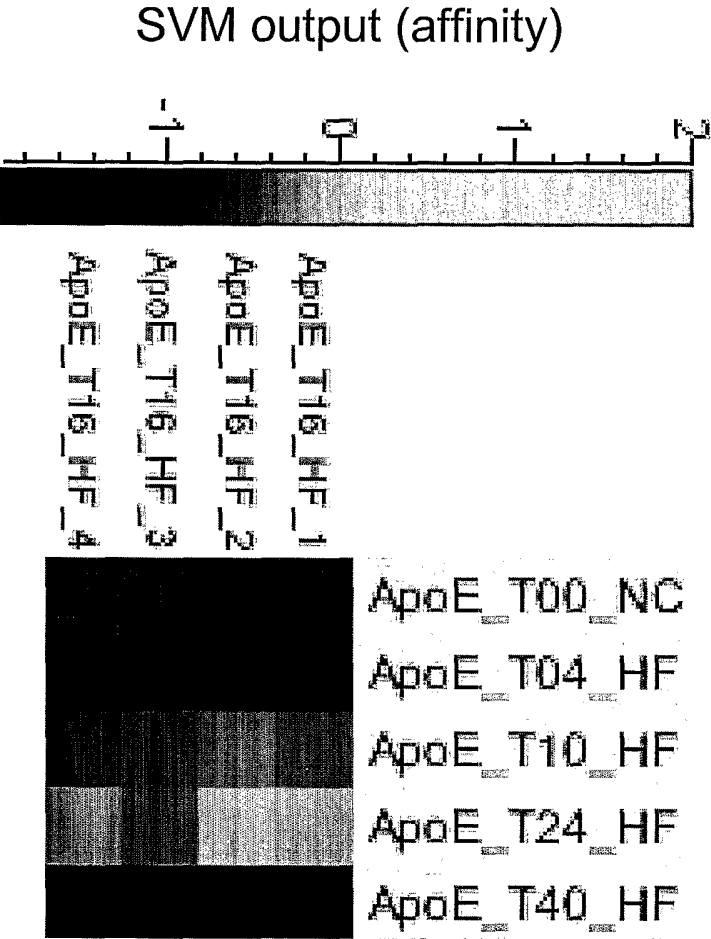


Figure 5B



Figure 7

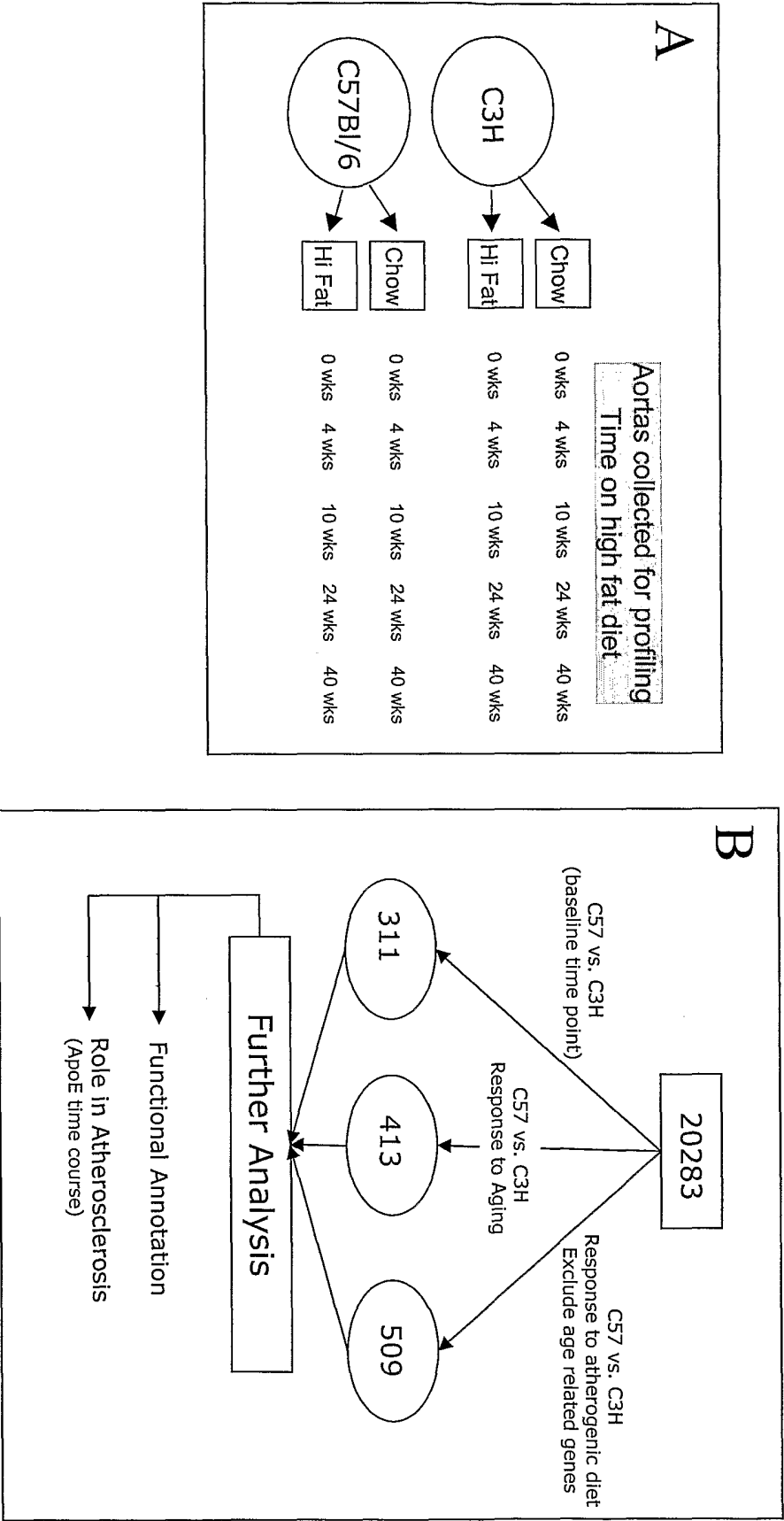
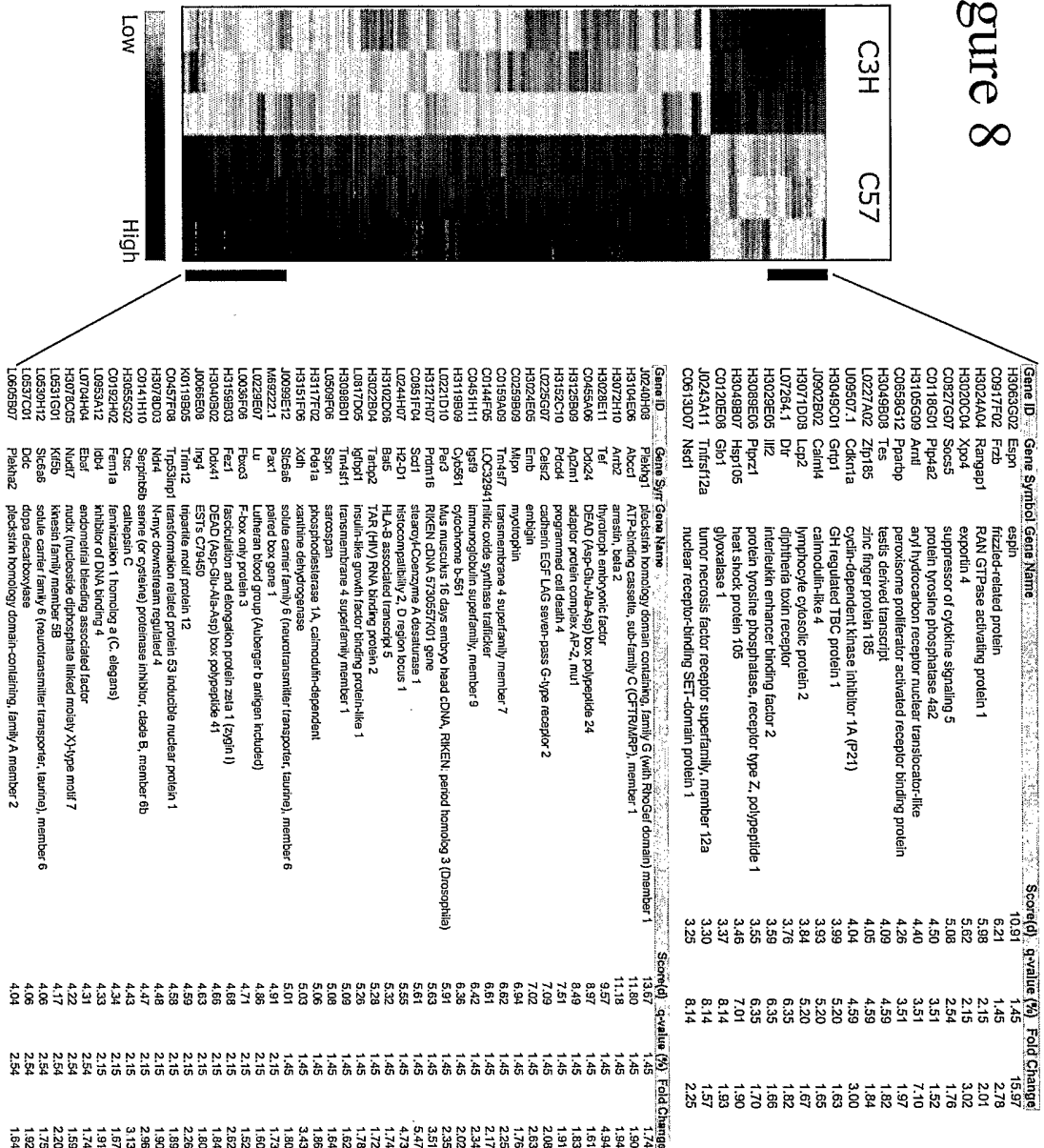


Figure 8



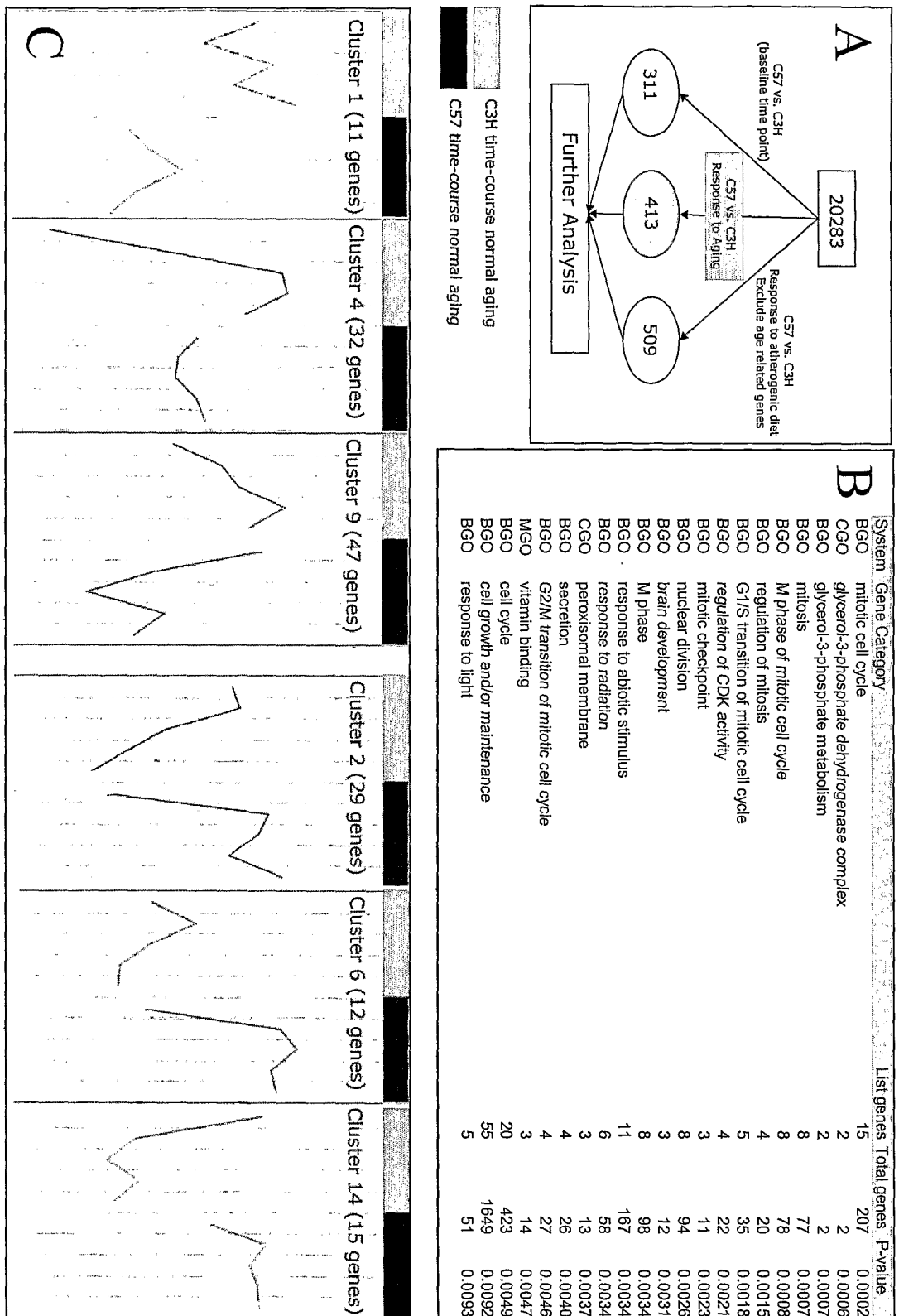


Figure 9

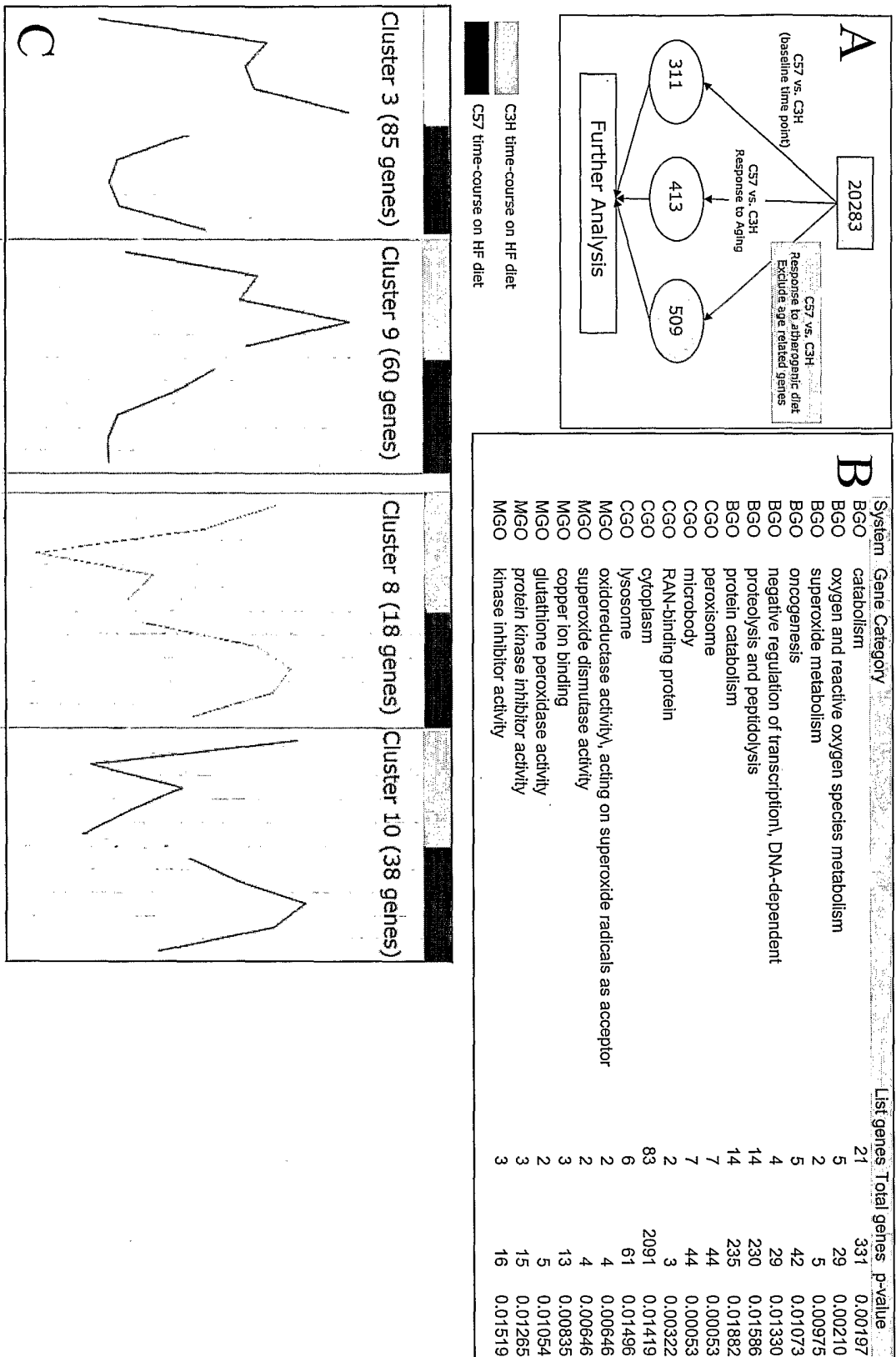
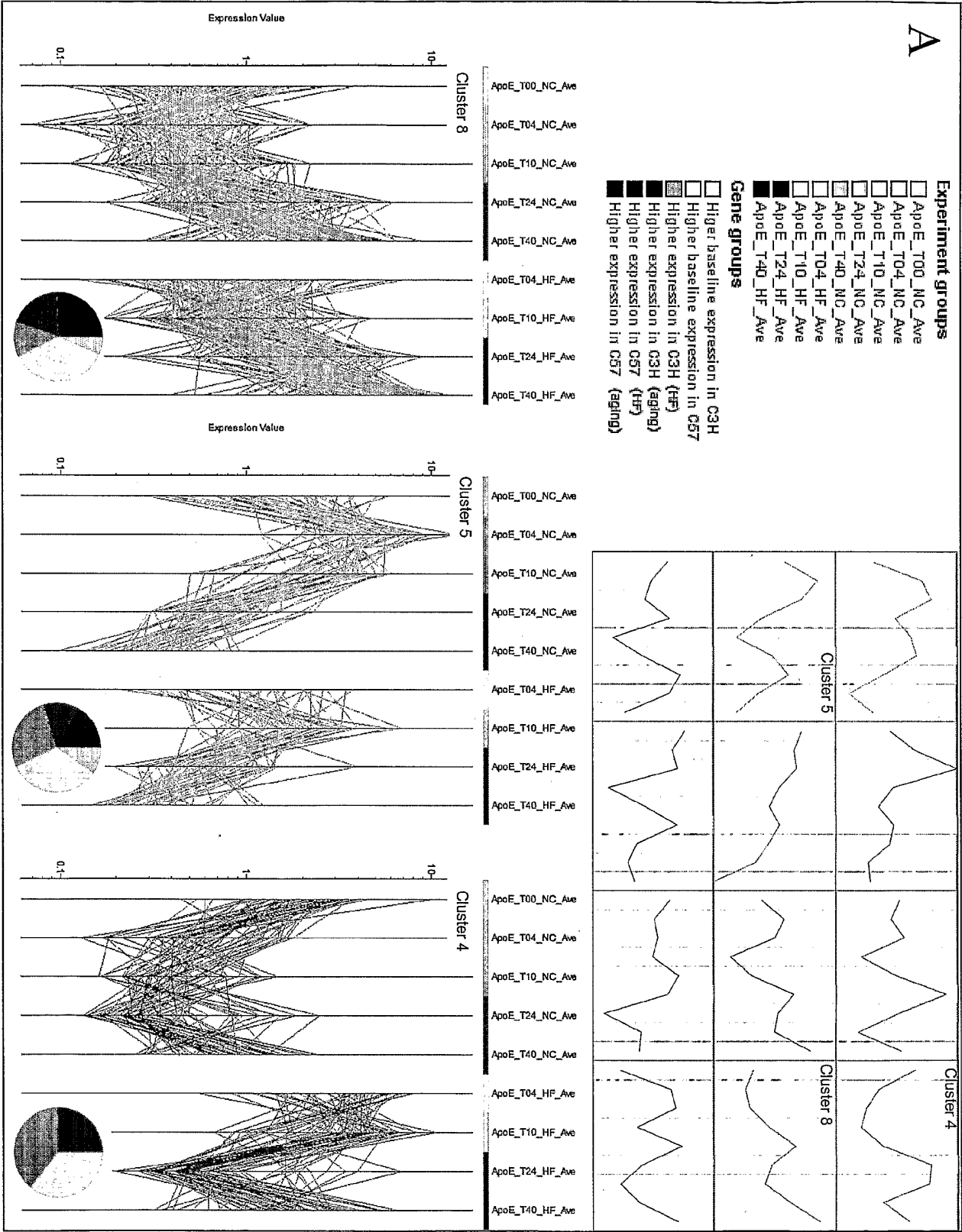


Figure 10

Figure 11A



Cluster 8 genes with higher expression in C57 at baseline

| Clone ID | Gene Symbol | Gene Name                                                                        |
|----------|-------------|----------------------------------------------------------------------------------|
| L0940E04 | Angp2       | angiopoietin 2                                                                   |
| C0175A04 | Anp32e      | acidic (leucine-rich) nuclear phosphoprotein 32 family, member E                 |
| L0503A06 | Ccd47       | CD47 antigen (Rb-related antigen, integrin-associated signal transducer)         |
| L0210E01 | Cln6        | ceroid-lipofuscinosis, neuronal 6                                                |
| H3055G02 | Ctsc        | cathepsin C                                                                      |
| H3117F10 | Ephx1       | epoxide hydrolase 1, microsomal                                                  |
| C0920F11 | Fthsdct1    | formyltetrahydrofolate synthetase domain containing 1                            |
| L0806E05 | Gtl2        | GTL2, imprinted maternally expressed untranslated mRNA                           |
| L0244H07 | H2-D1       | histocompatibility 2, D region locus 1                                           |
| H3134C05 | Mglap       | matrix gamma-carboxyglutamate (gla) protein                                      |
| M84918   | Myod1       | myogenic differentiation 1                                                       |
| H3117F02 | Pde1a       | phosphodiesterase 1A, calmodulin-dependent                                       |
| H3122D03 | Pdgfr       | platelet-derived growth factor, C polypeptide                                    |
| H3126B08 | Plazg7      | phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma) |
| C0141H10 | Serpinb6b   | serine (or cysteine) proteinase inhibitor, clade B, member 6b                    |
| J0099E12 | Sic6a6      | solute carrier family 6 (neurotransmitter transporter, taurine), member 6        |
| H3047C10 | Sic6a6      | solute carrier family 6 (neurotransmitter transporter, taurine), member 6        |
| L0530H12 | Sic6a6      | solute carrier family 6 (neurotransmitter transporter, taurine), member 6        |
| H3053B07 | Th          | tyrosine hydroxylase                                                             |
| C0159A09 | Tms4f7      | transmembrane 4 superfamily member 7                                             |
| K0179B05 | Trim12      | tripartite motif protein 12                                                      |
| H3151F06 | Xdh         | xanthine dehydrogenase                                                           |
| K0615F02 |             | RIKEN cDNA 1110056N09 gene                                                       |
| L0847B10 |             | RIKEN cDNA 2010106G01Rik gene                                                    |
| H3110A09 |             | RIKEN cDNA 2210408F21Rik gene                                                    |
| C0228C02 |             | RIKEN cDNA 2510004L01Rik gene                                                    |
| H3088G06 |             | RIKEN cDNA 3300001M20Rik gene                                                    |
| K0648G07 |             | RIKEN cDNA 9130002C22Rik gene                                                    |
| H3005C09 | A1325464    | expressed sequence A1325464                                                      |
| H3049F01 | A1427100    | expressed sequence A1427100                                                      |
| L0069H11 | AW550831    | ESTs AW550831                                                                    |
| H3127G02 | BG073838    | ESTs BG073838                                                                    |
| H3063C08 | BM241898    | ESTs BM241898                                                                    |
| J0066E08 | C79450      | ESTs C79450                                                                      |

Cluster 4 genes with higher expression in C3H on high fat diet

| Clone ID | Gene Symbol   | Gene Name                                                                          |
|----------|---------------|------------------------------------------------------------------------------------|
| J0211E11 | Ephx2         | epoxide hydrolase 2, cytoplasmic                                                   |
| L0521C08 | Mlycd         | malonyl-CoA decarboxylase                                                          |
| H3017C07 | Pdtx          | pyruvate dehydrogenase complex, component X                                        |
| H3153H05 | Peci          | peroxisomal delta3, delta2-enoyl-Coenzyme A isomerase                              |
| L0912B09 | Ppara         | peroxisome proliferator activated receptor alpha                                   |
| H3031D04 | Pp4a1         | protein tyrosine phosphatase 4a1                                                   |
| L0816B03 | Sirt3         | sirtuin 3 (silent mating type information regulation 2, homolog) 3 (S. cerevisiae) |
| L0054H07 | Slc2a4        | solute carrier family 2 (facilitated glucose transporter), member 4                |
| L0857H08 |               | RIKEN cDNA 1300003G02Rik gene                                                      |
| H3062H08 |               | RIKEN cDNA 1810030M08Rik gene                                                      |
| L0520B12 |               | RIKEN cDNA 2210010N04Rik gene                                                      |
| C0117D12 |               | RIKEN cDNA 2210016F16Rik gene                                                      |
| C0111F04 |               | RIKEN cDNA 3010027G13Rik gene                                                      |
| C0656F11 |               | RIKEN cDNA 4731402B04Rik gene                                                      |
| H3157C06 | B230217J03Rik | RIKEN cDNA B230217J03 gene                                                         |
| H3109E04 | BG072347      | ESTs BG072347                                                                      |
| G0102H10 | BI076467      | ESTs BI076467                                                                      |
| L0532G04 | BM123924      | ESTs BM123924                                                                      |
| L0544B11 | BM124723      | EST BM124723                                                                       |
| H3075H03 | D11Erd736e    | DNA segment, Chr 11, ERA TO Doi 736, expressed                                     |
| H3074D07 | MGC18837      | hypothetical protein MGC18837                                                      |
| C0811D11 |               | Mus musculus cDNA clone MGC:32180 IMAGE:5005221, complete cds                      |

Cluster 5 genes with higher expression in C3H at baseline

| Clone ID | Gene Symbol   | Gene Name                                                                  |
|----------|---------------|----------------------------------------------------------------------------|
| C0658G12 | Pparbp        | peroxisome proliferator activated receptor binding protein (pdx2 patirway) |
| H3029E05 | Ilf2          | interleukin enhancer binding factor 2                                      |
| L07264   | Dtr           | diphtheria toxin receptor (NF-AT pathway)                                  |
| J0020F08 | C76824        | expressed sequence C76824                                                  |
| C0173G07 | AW543759      | ESTs AW543759                                                              |
| H3136E03 | 5730444A13Rik | RIKEN cDNA 5730444A13 gene                                                 |
| H3034A06 | 11100180O8Rik | RIKEN cDNA 11100180O8 gene                                                 |

Figure 11B